이학박사학위논문

남극해 세균의 다양성 및 생태적 기능에 관한 연구

Study on the diversity and ecological functions of bacterial communities in the Southern Ocean

2017년 8월

서울대학교대학원
생명과학부
이 영 미

Study on the diversity and ecological functions of bacterial communities in the Southern Ocean

by Yung Mi Lee

Advisor: Professor Jongsik Chun, Ph.D

A Thesis Submitted for the Partial Fulfillment of the Degree of Doctor of Philosophy

August 2017

School of Biological Sciences

Seoul National University

ABSTRACT

The prokaryotes play significant roles in the biogeochemical cycles and remineralization of organic materials in the marine ecosystem. The Southern Ocean (SO) is of global significance because of its role against global warming through the absorption of carbon dioxide and formation of cold Antarctic bottom water. Since ecological functions of microorganisms including carbon fixation and nutrient cycling in SO are directly linked to the global ecosystem, understanding or prediction of the microbial functions in SO is important. However, few studies regarding diversity and ecological functions of bacteria in SO have been performed until now mainly due to the difficulties in access to this area. In addition, most studies on the bacterial diversity were performed by traditional molecular approaches such as fingerprinting and clonal sequencing of 16S rRNA gene and it led to limited understanding of the full extent of diversity.

In this study, to expand the knowledge on the bacterial diversity of seawater and sediments in the SO, high-throughput sequencing of 16S rRNA gene was applied. Moreover, since the patterns of bacterial diversity are largely determined by physicochemical parameters, the relationship between the bacterial diversity and environmental factors was explored to provide insights into the change pattern of each bacterial group. In addition, as a basis of physiological studies, bacterial culture from marine sediments was performed and the physiological characteristics of culturable bacteria were investigated. Finally, to understand the metabolic potential and ecological functions of JS1 lineage of *Atribacteria*, a predominant but uncultivated bacterial group in anoxic marine sediment, single cell genomics was applied to analyze the genomic features.

Bacterial diversity of seawaters and the sediments was analyzed according to regional and local scale. On the regional scale study, bacterial community structure of the surface seawaters across 1600 to 3000 km in SO according to the frontal systems was investigated. On the local scale study, the vertical and horizontal variability of bacterial

communities in the seawater and sediments of the Ross Sea was performed. Characterization of the bacterial communities at different water depths was performed at greater resolution covering various regions over the Terra Nova Bay (TNB) of the Ross Sea. In addition, the dynamics of the bacterial community in surface waters was carried out to investigate whether the surface water bacterial community is affected by the local conditions such as distance from glacier front or land. The benthic bacterial diversity, along with the spatial variability in the Ross Sea including sediments from Australian-Antarctic Ridge was analyzed.

The regional study revealed that three bacterial groups, *Bacteroidetes*, *Alphaproteobacteria*, and *Gammaproteobacteria* dominated in seawaters in the SO. Regardless of the dominance of these three groups, relative abundance of *Cyanobacteria* and *Actinobacteria* diminished or disappeared in the south of Subantarctic Front and *Verrucomicrobia* and SAR406 was not recovered in south of Polar Front indicating the distribution of specific group is influenced differently by fronts. The clear biogeographic pattern in the bacterial communities according to the frontal systems was revealed implying the different and distinct metabolic traits and functions exist in each frontal zone. In addition, bacterial communities were clearly divided by the Subantarctic Front and Southern Boundary, the northernmost and southernmost boundaries of the Antarctic Circumpolar Current (ACC), respectively, showing significant correlations with environmental variables. Among the physicochemical variables with strong correlation each other, temperature best determined the community separation.

The local scale study on the bacterial diversity in the seawater of the Ross Sea revealed highly diverse bacterial lineages which were not recovered using traditional molecular approaches in this region. The bacterial community composition in the seawater of the Ross Sea according to the vertical and horizontal variability showed the clear vertical stratification of bacterial diversity along the depth and less differentiation according to the

horizontal distance. Three bacterial groups, *Bacteroidetes*, *Alphaproteobacteria*, and *Gammaproteobacteria*, dominated throughout the whole water column. Along the depth increase, *Deltaproteobacteria*, SAR406, *Verrucomicrobia*, *Planctomycetes*, *Lentisphaerae*, *Chloroflexi*, *Gemmatimonadetes*, *Planctomycetes*, and *Firmicutes* which were rarely recovered in surface waters, increased and the bacterial diversity also increased. Horizontal distribution of bacterial composition revealed the high portion of *Bacteroidetes* near the land indicating the association of this group with algal bloom. The vertical and horizontal variability of bacterial communities indicated that the steep gradient in the environments along the depth than the horizontal distance is closely related with bacterial assemblage variability. Among the physicochemical parameters, oxygen level and concentration of nitrite and nitrate explained well about this variability.

The spatial comparison in the bacterial communities of marine sediment revealed diverse bacterial populations compared to previous studies in this region and the presence of the high portion of uncultivated groups. Distinct bacterial communities between aerobic and anaerobic sediments outcompeting the geological location were revealed. In the aerobic surface sediments, *Gammaproteobacteria*, *Deltaproteobacteria*, *Bacteroidetes*, and *Planctomycetes* dominated while Ca. *Atribacteria*, *Chloroflexi*, and *Actinobacteria* were found in higher abundance in the anoxic sediment horizon. Also the portion of candidate phyla was higher in anoxic sediments than surface samples. This indicated the redox state in the sediments plays significant roles for shaping the bacterial community structures.

To further examine the ecological functions of bacteria, marine sediments with a large portion of uncultivated groups were cultivated in oligotrophic medium at low temperature. In spite of the limited coverage of cultivated bacterial diversity, many candidate novel species with less than 98.65% 16S rRNA gene similarity to known species were recovered. Forty-six and 25% of bacterial isolates showed extracellular protease

and/or lipase activities indicating that the isolated strains contribute to the hydrolysis of major organic constituents and are therefore involved in carbon and nitrogen cycling at the low temperature of sediments of Ross Sea.

The ecological function of JS1 lineage of candidate phylum *Atribacteria*, which was known to be predominant in the anoxic sediments, has been poorly understood due to the lack of cultivated bacterial strains until now. Since the high portion of Ca. *Atribacteria* JS1 lineage in the anoxic sediments of the Ross Sea was revealed through the bacterial community investigation, single cell sorting and amplification of genomic DNA as complements for cultivation method were applied to elucidate the metabolism and function of this group. Eighteen single cell amplified genomes of JS1 species have been obtained and co-assembled. The highest coverage JS1 genome revealed how this species is dominant in anaerobic environments of the Ross Sea. The metabolic versatility of this species as an acetate producer and syntrophic acetate oxidizer seemed to show increased survivability in nature through syntrophic interactions with partner methanogens or using various heterotrophic substrates to outcompete other primary and secondary fermenters. This versatile growth mode seemed to explain the dominance of Ca. *Atribacteria* JS1 within anoxic sediments of the Ross Sea. Thus, the genomic insights from this study shed light on the ecological function of Ca. *Atribacteria* JS1 in methane-related carbon cycling.

Keywords: marine bacteria, bacterial diversity, pyrosequencing, cultivation, single cell genomics, exoenzymatic profiles, candidate phylum, Atribacteria JS1, Ross Sea, Southern Ocean

ABBREVIATIONS

SO: Southern Ocean

STF: Subtropical front

SAF: Subantarctic front

PF: Polar front

SACCF: Southern antarctic circumpolar current front

ACC: Antarctic circumpolar current

SAZ: Subantarctic zone

PFZ: Polar frontal zone

SACCZ: Southern antarctic circumpolar current frontal zone

AZ: Antarctic zone

CDW: Circumpolar deep water

AABW: Antarctic bottom water

AAIW: Antarctic intermediate water

SAMW: Subantarctic mode water

TNB: Terra Nova Bay

NGS: Next generation sequencing

OTU: Operational taxonomic unit

Chl a: Chlorophyll a

rRNA: ribosomal ribonucleic acid

NMDS: Non-metric multidimensional scaling

PCA: Principal component analysis

ANOSIM: Analysis of similarity

RDP: Ribosomal database project

PCR: Polymerase chain reaction

MDA: Multiple displacement amplification

ANI: Average nucleotide identity

DGGE: Denaturing gradient gel electrophoresis

T-RFLP: Terminal restriction fragment length polymorphism

Ca.: Candidatus

SAG: Single-cell amplified genome

cSAG: composite SAG

EC: Enzyme commission

BLAST: Basic local alignment search tool

KEGG: Kyoto encyclopedia of genes and genomes

KAAS: KEGG automatic annotation server

CSCG: Conserved single copy gene

ABC transporter: ATP-binding cassette transporter

BMC: Bacterial microcompartments

SAOB: Syntrophic acetate-oxidizing bacterium

TC: Total carbon

TOC: Total organic carbon

TIC: Total inorganic carbon

TN: Total nitrogen

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1. Introduction: Oceanography and microbial ecology of the Southern Ocean and Ross Sea

1.1 Oceanography of the Southern Ocean and Ross Sea

The Southern Ocean (SO) encircles Antarctica, and is generally accepted to be south of 60° S latitude (Fig. 1-1a), and comprises the southernmost waters of the world ocean (Pakhomov et al. 2000). Being the fourth-largest of the five oceanic divisions, the SO plays a significant role against global warming due to its buffering capacity, which allows it to absorb about 40% of the anthropogenic carbon dioxide released into the atmosphere and facilitates heat transport (Mikaloff Fletcher et al. 2006). The SO acts as an interface between the deep ocean and the atmosphere through the wind-driven strong upwelling that brings carbon and nutrient-rich deep water to the surface where gases can begin to equilibrate with the atmosphere (Hauck et al. 2015). Thus, changes such as reduction in the Antarctic ice sheet and the capacity of the SO to transport heat and absorb atmospheric CO_2 significantly affect global climate (Cavicchioli 2015). In addition, the oceanic expanse of the SO is large enough that the carbon fixation and nutrient cycling occurring there are directly linked to the global ecosystem (Cavicchioli 2015, Wilkins et al. 2012).

The SO is composed of several water masses that are physicochemically distinct but linked by circulation (Talley et al. 2011, Wilkins et al. 2013). Highly saline Lower Circumpolar Deep Water (CDW) originating in the Atlantic and oxygen-depleted CDW from the deep Indian and Pacific oceans flow poleward and shoal to reach the sea surface of the Antarctic Zone (Wilkins et al. 2013) (Fig. 1-1b). Some of the upwelled CDW that moves southward becomes colder and denser and then sinks to form Antarctic Bottom Water (AABW) (Wilkins et al. 2013). Northward spreading AABW ventilates the densest layers of the global ocean. The remainder of the upwelled CDW becomes warmer and freshened as it moves northward before then sinking to form Antarctic Intermediate Water (AAIW), which spreads northward at about a 1,000 m depth. Overlying the AAIW is the warmer Subantarctic Mode Water (SAMW) (Wilkins et al. 2013) (Fig. 1-1b). The SO is

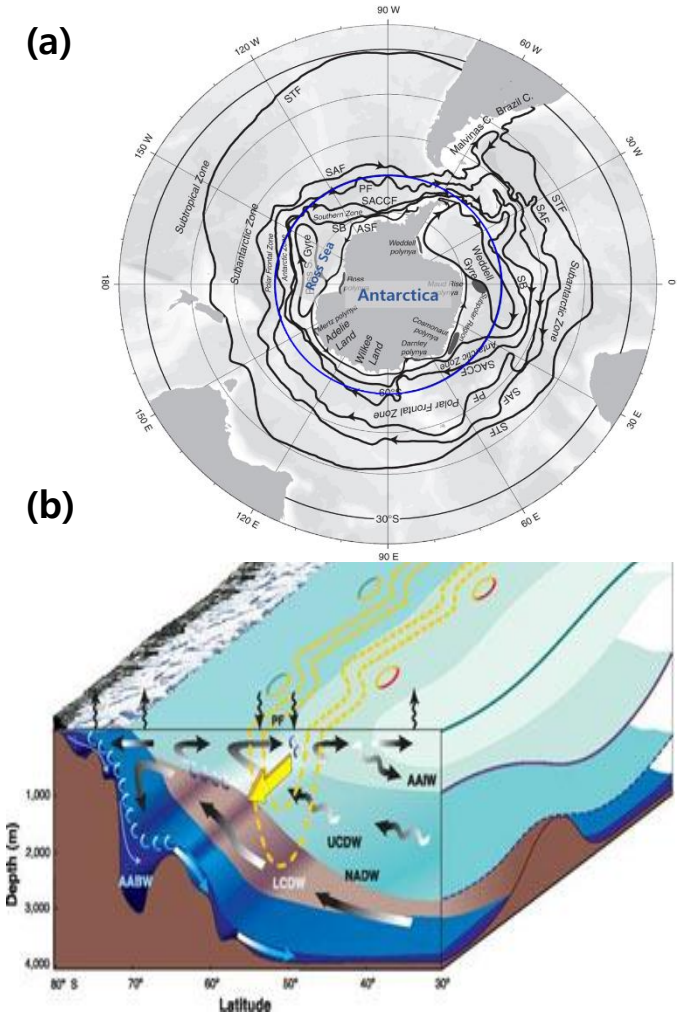


Figure 1-1. A schematic view of the Southern Ocean and principal fronts and oceanographic zones in the Southern Ocean (a), and meridional overturning circulation (b). Abbreviations are as follows: AAIW: Antarctic Intermediate Water, AABW: Antarctic Bottom Water, NADW: North Atlantic Deep Water, LCDW: Lower Circumpolar Deep Water, UCDW: Upper Circumpolar Deep Water, SAF: Subantarctic Front, PF: Polar Front, SACCF: Southern Antarctic Circumpolar Current Front, SB: Southern Boundary, ACC: Antarctic Circumpolar Current, [Images adapted from Talley et al., (2011) (a) and Woods Hole Oceanographic (b)]

subdivided into several distinct zones and is not ecologically uniform due to the circumpolar fronts such as the Subtropical Front (STF), Subantarctic Front (SAF), Polar Front (PF), and Southern Antarctic Circumpolar Current Front (SACCF) along the latitudinal distribution (Deacon 1982, Ghiglione et al. 2012, Orsi et al. 1995, Pakhomov and McQuaid 1996, Talley et al. 2011, Wilkins et al. 2012) (Fig. 1-1a). The Antarctic Circumpolar Current (ACC), composed of the SAF, PF, and the SACCF is the largest and strongest wind-driven eastward flowing current on Earth; isolating Antarctica from other oceans and is also the only ocean current to travel all the way around the planet (Orsi et al. 1995).

The Ross Sea is a large embayment on the Antarctic coastline (Fig. 1-1a), covering an area of approximately 750,000 km^2, with an average water depth of 500 m (Faranda et al. 2012). The development of the Ross Sea polynya, the largest consistently forming polynya on Earth, is due to the physiography of the terrestrial region where strong katabatic winds flow from the land over the ocean and continually blow away sea ice (Arrigo and van Dijken 2004). Due to the polynyas, the Ross Sea experiences one of the most spatially expansive phytoplankton blooms in the SO and as such, is one of the most productive Antarctic shelf systems (Arrigo et al. 1998a, Arrigo et al. 1998b). With seasonally high primary production, the organic carbon in the Ross Sea is likely to be of marine origin, and is mainly produced by autotrophic eukaryotes (primarily diatoms). The particle flux may enable the entire particulate pool to accumulate on the seafloor within 2-3 days, making the Ross Sea a repository for larger volumes of sediment (Albertelli et al. 1998, Ship et al. 1999).

1.2 Microbial ecology in the Southern Ocean

1.2.1 Studies on bacterial diversity in the Southern Ocean

Prokaryotes with substantial amounts of biomass play a significant role in the biogeochemical cycles and remineralization of organic materials in the marine ecosystem (Whitman et al. 1998). Investigating the microbial diversity and the determinant that shapes microbial structure is an important first step in understanding the ecological roles of microorganisms. In this context, an investigation of the global scale pattern of bacterial diversity in oceans encompassing large environmental gradients has been performed in order to provide fundamental answers to long-standing questions in the field of marine microbial ecology (Fuhrman et al. 2008, Lozupone and Knight 2007, Sunagawa et al. 2015, Zinger et al. 2011). The global scale study identified temperature as an important factor in shaping the bacterial diversity found in seawater (Fuhrman et al. 2008, Lozupone and Knight 2007, Sunagawa et al. 2015, Zinger et al. 2011). However, since light penetration, availability of labile organic matter, and salinity have also been reported as key factors affecting bacterial diversity in different study areas (D/'Hondt et al. 2015, Glud 2008, Herlemann et al. 2011), determining the environmental factor that explains the bacterial diversity pattern of the SO can provide significant insights for the better understanding or prediction of the community function in response to environmental changes.

Microorganisms in the SO are recognized as having crucial roles in the global biogeochemical cycle, including carbon sequestration (Cavicchioli 2015). Bacterial growth rates in cold polar seas, which are lower than those in warmer oceanic systems, increase faster with rising temperatures, indicating that microbial processes in polar waters are particularly sensitive to small changes in the environment and that changes will potentially have an influence on ecosystem function (Kirchman et al. 2009). Thus, understanding the structure and function of microbial communities across fronts with contemporaneous physicochemical differences in the SO may provide insights into the

response of polar microorganisms to upcoming environmental changes. In this context, studies on the biogeographic distribution of bacteria in the SO according to separate frontal zones with different hydrological properties have been performed (Table 1-1) (Abell and Bowman 2005, Giebel et al. 2009, Pakhomov and McQuaid 1996, Selje et al. 2004, Ward et al. 2003, Wilkins et al. 2012). However, identifying the biogeographical boundary that separates the distribution of microbial composition in the SO is still problematic. In most studies, the PF was suggested to be the biogeographic barrier (Abell and Bowman 2005, Giebel et al. 2009, Pakhomov and McQuaid 1996, Selje et al. 2004, Wilkins et al. 2012). The biogeographical importance of the SAF to the distribution of species was less obvious, but in some studies the SAF was suggested to be an important faunistic boundary (Deacon 1982, Froneman et al. 1995, Pakhomov et al. 1999, Pakhomov et al. 2000).

Few studies of the bacterial diversity in the Ross Sea have been performed by comparison to the intensively performed studies of the hydrological dynamics and phytoplankton assemblages (Faranda et al. 2012). Considering that the ocean is subdivided into the pelagic and benthic zones, which are fundamentally different in terms of their physical, chemical, and biological properties (Austen et al. 2002, Whitman et al. 1998, Zinger et al. 2011), studies on bacterial diversity in the Ross Sea have been investigated separately at these two zones (Table 1-1). In seawater, studies of bacterial diversity have been investigated with regard to human impact, water mass, depth, sea-ice melting or nutrient availability using molecular approaches such as 16S rRNA gene cloning, fingerprinting methods, or cultivation associated with fluorescent *in situ* hybridization technique (Table 1-1) (Celussi et al. 2009a, Celussi et al. 2009b, Celussi et al. 2010, Fabiano et al. 1995, Fabiano et al. 2000, Gentile et al. 2006, Lo Giudice et al. 2012, Michaud et al. 2004, Povero et al. 2006, Yakimov et al. 2004).

Similar to continental shelf systems, polar shelves are regions of intense biological activity and biogeochemical cycling (Ducklow et al. 2008, Fonda Umani et al.

2005). Seasonal high primary production and intense sedimentation rates may fuel the benthic microbial communities by pulsed organic inputs to the sediments during the spring–summer period in the Ross Sea. In this context, the relationship between labile organic matter and the bacterial communities in the shallow sediments have been analyzed using 16S rRNA gene cloning and fingerprinting methods (Baldi et al. 2010). The microbial diversity was also analyzed using pyrosequencing both in the surface sediments and in the deep-sea sediments that were deposited below the permanent Ross Sea Ice Shelf, the world's largest ice shelf which covers 560,000 km^2 (Carr et al. 2013, Learman et al. 2016).

In summary, understanding the bacterial diversity in the SO is important since microbial process in the SO can influence global ecosystems. Thus, studies of the bacterial diversity in relation to environmental parameters have been performed in the SO to consider the global impact. However, most of the studies have been performed using traditional molecular approaches such as 16S rRNA gene cloning, fingerprinting such as denaturing gradient gel electrophoresis (DGGE), and terminal restriction fragment length polymorphism (T-RFLP) with low resolution in order to unveil the full extent of bacterial diversity. Thus, studies using next generation sequencing (NGS) are required to expand the knowledge base on bacterial diversity present in seawater and marine sediments of the SO, to help determine and understand the relationship between bacterial diversity and environmental parameters.

Table 1-1. Summary of previous studies on bacterial diversity in the Southern Ocean and Ross Sea

Location	Habitat	Main interests	Method	Reference
Southern Ocean	Seawater	Distribution of Roseobacter along the circumpolar front	FISH	Selje et al. (2004)
		Distribution of Flavobacteria along the circumpolar front	FISH and DGGE of 16S rRNA gene	Abell and Bowman (2005)
		Bacterial diversity along the latitude	Enzymatic activities, T-RFLP, and cloning of 16S rRNA gene	Baldwin et al. (2005)
		Distribution of Roseobacter RCA and SAR11 along the circumpolar front	DGGE of 16S rRNA gene	Giebel et al. (2009)
		Bacterial diversity along the latitudinal change	Catalyzed reporter deposition–fluorescence in situ hybridization	Weitz et al. (2010)
		Pole-to-pole biogeography in bacterial diversity	Pyrosequencing of 16S rRNA gene	Ghiglione et al. (2012)
		Bacterial diversity and function by the Polar Front	Pyrosequencing of metagenomic DNA	Wilkins et al. (2012)
		Advection and bacterial diversity	Pyrosequencing of 16S rRNA gene	Wilkins et al. (2013)
		Bacterial diversity in polar region	Pyrosequencing of 16S rRNA gene	Sul et al. (2013)
		Global pattern of bacterial diversity	Illumina sequencing of metagenomic	Sunagawa et al. (2015)
	Seawater and sediment	Global pattern of bacterial diversity	Pyrosequencing of 16S rRNA gene	Zinger et al. (2011)
Ross Sea	Seawater	Effect of oil on bacterial diversity from surface water	T-RFLP and cloning of 16S rRNA gene	Yakimov et al. (2004)
		Diversity and physiological characteristics of culturable bacteria from surface water	Cultivation and physiological characteristics determination	Michaud et al. (2004)

Table 1-1. To be continued

Location	Habitat	Main interests	Method	Reference
Ross Sea	Seawater	Bacterial diversity of surface water	Cloning of 16S rRNA gene	Gentile et al. (2006)
		Bacterial diversity along the depth	DGGE of 16S rRNA gene	Celussi et al. (2009)
		Bacterial diversity according to water mass	DGGE of 16S rRNA gene and enzyme	Celussi et al. (2010)
		Diversity and physiological characteristics of culturable bacteria	FISH, cultivation, and amplified rDNA	Lo Giudice et al. (2012)
	Sediment	Enzymatic activity and bacterial distribution according to the organic	Bacterial counting and enzymatic	Fabiano et al. (1998)
		Bacterial diversity of surface sediment	T-RFLP and cloning of 16S rRNA gene	Baldi et al. (2010)
		Bacterial diversity along the depth	Pyrosequencing of 16S rRNA gene	Carr et al. (2013)
		Diversity and physiological characteristics of culturable bacteria from	Cultivation and enzyme activity	Lee et al. (2014)
		Bacterial diversity of surface sediment	Illumina MiSeq sequencing of 16S	Learman et al. (2016)
	Seawater and sediment	Diversity and physiological characteristics of culturable bacteria	Cultivation and enzyme activity	Giudice et al. (2006)

1.2.2 Methodologies for studies of bacterial ecological function

NGS technologies have revolutionized the field of microbial ecology by overcoming very low detection limits, which allows researchers to reach a nearly complete sample characterization for microbial diversity. However, in spite of the enormous progress in the investigation of bacterial diversity with the advancement of this sequencing technology, the community structure cannot directly be used to obtain the ecological functions that the microorganisms perform in nature. Instead, investigating the physiological characteristics, activities, and metabolisms of the microorganisms is required to provide answers to questions about their *in situ* ecological functions. Thus, elucidating the metabolic activity and ecological function of marine bacteria is the challenging next step to perform after understanding the bacterial diversity in microbial ecology studies.

To examine the physiological characteristics and ecological functions of the representative bacteria, cultivation methods provide the most fundamental approach (Fig. 1-2). In spite of the fact that approximately less than 1% of the total bacterial groups in nature are amenable to being cultivated (Amann et al. 1995), bacterial isolates make it possible to investigate the metabolic traits and validate hypotheses regarding the ecological functions of at least some members of the microbial community. In order to overcome the limitations of cultivation for determining the ecological function of bacteria, exoenzymatic profiles such as proteinase, lipase, ß-glucosidase, and alkaline phosphatase can be used as indicators for the presence of metabolically active bacteria using fluorescently labeled substrate analogues in natural samples (Celussi et al. 2009a, Celussi et al. 2010, Coolen and Overmann 2000).

Metabolic reconstruction from genomic analyses provides information on the metabolism potential and putative ecological functions of a microbial species. The advancement of sequencing technology enables researchers to recover draft or nearly

complete genomes from environmental DNA and also minute amounts of DNA from a single cell (Nobu et al. 2015, Rinke et al. 2013). Through the use of metagenomics, a powerful tool to obtain the genomic information of microorganisms by direct sequencing of environmental DNA (Handelsman 2004), some draft genome sequences, including candidate phyla, have been recovered (Dodsworth et al. 2013, Hedlund et al. 2014, Hug et al. 2013, Nobu et al. 2015, Rinke et al. 2013). However, the high degree of genomic heterogeneity in natural populations makes it difficult to obtain sufficient information from previously uncultivated bacteria by metagenome analysis (Rinke et al. 2013). Whole-genome sequencing of minute amounts of DNA from a single cell has allowed us to obtain the genomes of uncultivated microorganisms regardless of heterogeneity expanding our knowledge on microbial taxonomic and functional diversity (Eloe-Fadrosh et al. 2016, Kamke et al. 2013, Kolinko et al. 2012, Macaulay and Voet 2014, Rinke et al. 2013).

DNA stable-isotope probing is a method in which the incorporation of a stable isotope from a labeled substrate is used to identify the function of microorganisms in the environment (Chen and Murrell 2010). Stable-isotope probing offers a powerful technique that can be used to identify microorganisms that are actively involved in specific metabolic processes under the types of conditions where those processes occur *in situ*. In addition, the coupling of DNA stable-isotope probing with metagenomics can link the identity of the bacteria with their function in the environment (Chen and Murrell 2010, Radajewski et al. 2000). Results from these investigations could provide insights needed to provide favorable cultivation conditions of uncultivated groups, which can be a useful resource for the evaluation of ecological functions in nature.

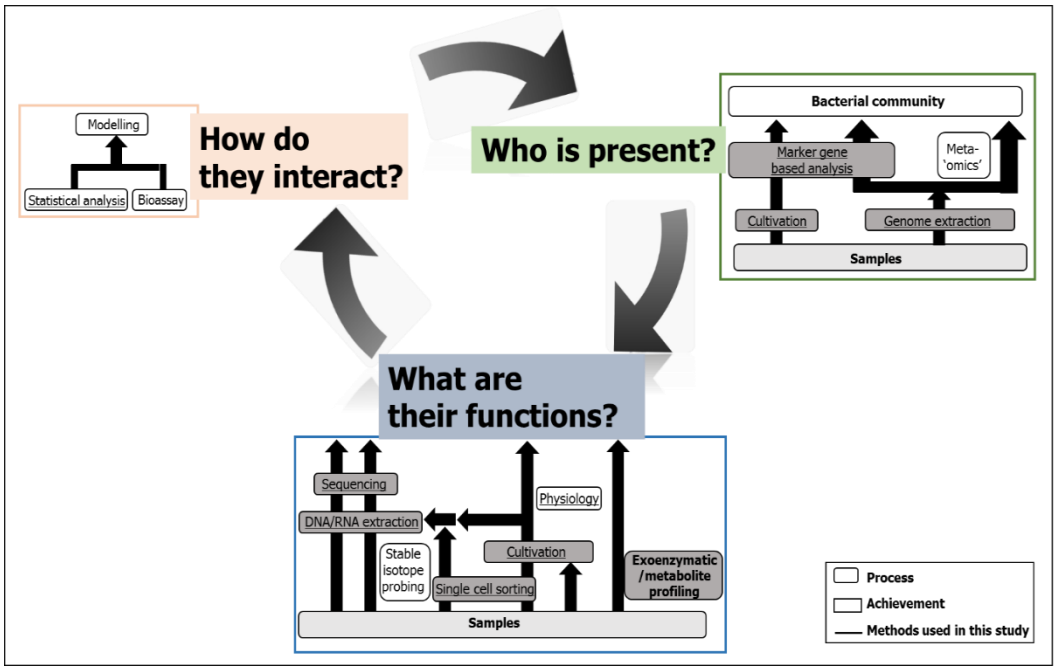


Figure 1-2. Concept of microbial ecology and methodologies used for investigating interests in microbial ecology

1.3 Objectives of this study

Considering the sensitivity of the SO bacteria to climate change, understanding the diversity and ecological functions of bacterial communities is of significant interest. Characterizing the bacterial diversity at deeper depths is an important preliminary step towards understanding roles of the bacteria in this important environment. In addition, since the patterns of bacterial diversity are largely determined by physicochemical parameters, exploring the relationship between the bacterial diversity and environmental factors can provide insights into the change pattern of each bacterial group. Further, to investigate the metabolic potential using both culture-dependent and -independent approaches can provide important information about the putative natural ecological function using the advantages that each method can provide. In this context, this study was aimed at understanding the ecological functions of bacteria in the SO based on detailed characterizations of the bacterial community structure (Fig. 1-3).

In chapter 2, bacterial diversity was investigated on the regional and local scale with the correlated environmental factors. On the regional scale, bacterial populations in the SO covering an area from 47°S to 74°S across circumpolar fronts, the distribution of taxonomic groups or species by the presence of each front, and the key determinant responsible for shaping the structure were investigated. On the local scale, the bacterial diversity of the seawater and sediments in the Ross Sea were investigated. Characterization of the bacterial communities at different water depths was performed at greater resolution covering various regions over the TNB of the Ross Sea. In addition, the dynamics of the bacterial community in surface waters was carried out to investigate whether the surface water bacterial community is dynamically affected by the local conditions such as distance from glacier front or land. The benthic bacterial diversity, along with the spatial variability in the Ross Sea, was analyzed. For better characterization of the composition of the bacterial community and the diversity of seawater and sediments, the 16S rRNA gene was sequenced using high-throughput 454 pyrosequencing.

Based on the diversity results, the metabolic potential and putative ecological

functions were investigated. In chapter 3, to gain insights into the culturable diversity and ecological roles of the bacteria, marine sediment samples from the Ross Sea which were selected based on the bacterial diversity results from the 16S rRNA gene pyrosequenicng were cultivated. The taxonomic affiliations of the bacterial isolates and the physiological characteristics such as growth response to the temperature and production of extracelluar enzymes were determined in order to infer the natural ecological functions. In chapter 4, genomes obtained from single-cell sorting and amplification were analyzed to determine the metabolic features and predict the ecological functions of the *Atribacteria* JS1 lineage with no cultured representatives, in spite of the predominance of the strain throughout the anoxic sediments of the Ross Sea.

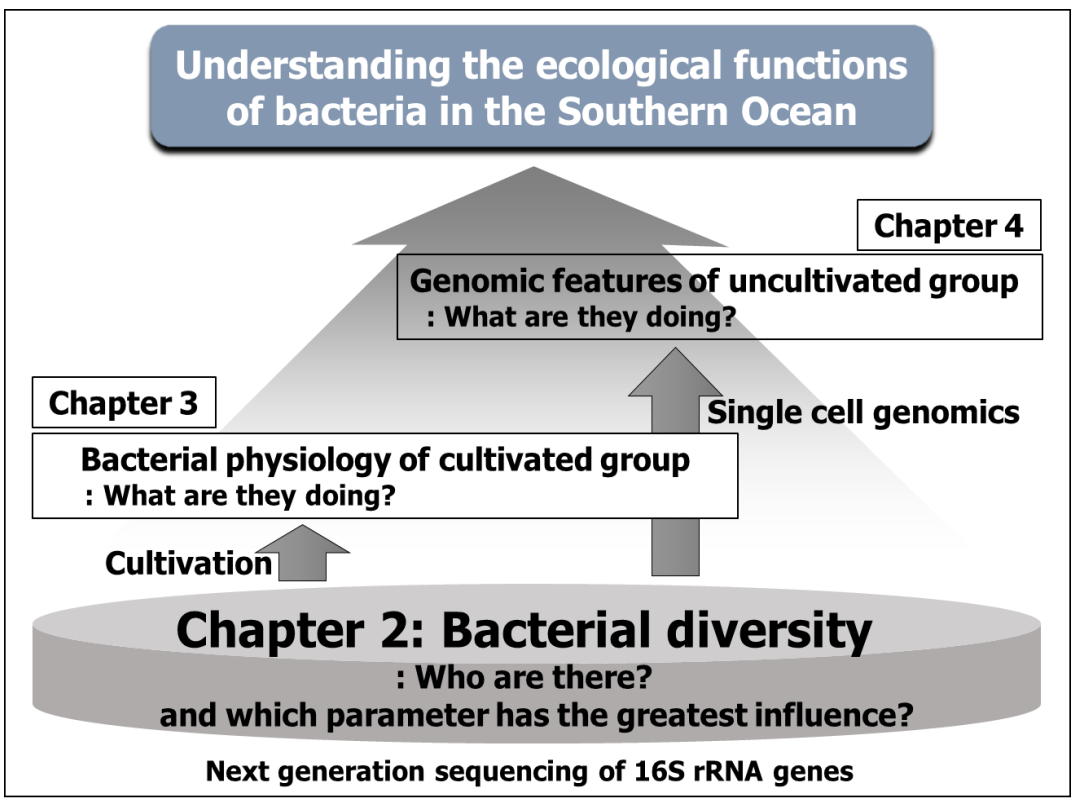


Figure 1-3. Schemes and purposes of this study

CHAPTER 2. Bacterial community structure in seawater and sediments of the Southern Ocean and the Ross Sea

2.1 Introduction

Microorganisms constitute dominant components of marine plankton and play significant roles in mediating a wide range of biogeochemical cycles in marine environments (Arı´stegui et al. 2009, Azam et al. 1983, Fuhrman et al. 1989, Zinger et al. 2011). Investigating the microbial diversity is an important first step for understanding the ecological roles of microorganisms. Microbial community structures are strongly affected by environmental factors. Therefore, understanding key factors that exert large effects on the microbial composition, in turn causing different ecological functions, is fundamental and important in microbial ecology. Considering the global significance of SO and its large expanse, ecosystem function changes in the SO in response to microbial community shifts occurring in parallel with environmental or climatic changes may affect global biogeochemical cycles (Cavicchioli 2015, Mikaloff Fletcher et al. 2006, Pakhomov et al. 2000, Wilkins et al. 2012). Since SO is not ecologically uniform and subdivided into several distinct zones with different physicochemical properties which are separated by circumpolar fronts, it plays as a natural laboratory to investigate the bacterial response over the coming environmental changes.

The Ross Sea forms unique ecosystems distinct from other areas of the SO because of the extreme seasonality, numerous polynyas, areas of open water surrounded by sea ice, and nearby extensive ice shelf (Arrigo et al. 1998a). In particular, TNB of the Ross Sea, bordered by the Drygalski Ice Tongue on the south and the Campbell Ice Tongue in the north usually shows a strongly stratified water column in the summer (Celussi et al. 2009a, Mangoni et al. 2004). Surface water is characterized by high exposure to solar radiation and a steep gradient in temperature and is heavily influenced by the local atmosphere and by sea ice freezing and melting processes in polar seas; in contrast, deep water is characterized by the absence of light and more quiescent conditions with consistent low temperature and high hydrostatic pressure (Celussi et al. 2009a, Qian et al. 2011). These contrasting conditions are likely to act as environmental filters to influence

the preferred or well-adapted bacterial assemblages at different depths within a water column with varying timescales, i.e., rapid speciation as a consequence of dynamic conditions at the surface and more time to select phenotypes for specific conditions in the deep sea (Ghiglione et al. 2012). Since depth-related bacterial communities reflect the metabolic traits that may determine the fate of organic matter degradation and utilization, understanding bacterial communities along the depth can provide biogeochemical roles of bacteria.

TNB of the Ross Sea is the regions of intense biological activity and biogeochemical cycling in relation to the algal bloom that resulted from the intense polynyas during spring-summer seasons (Baldi et al. 2010). In this area, organic carbon is considered to be marine origin and the particulate flux may affect the organic matter pool and heterotrophic bacteria community in the benthic ecosystem (Baldi et al. 2010). Bacterial diversity in relation to different available organic matter inputs or physiological characteristics has been investigated to provide insights toward knowledge of bacteria in Antarctic benthic ecosystem by culture-based or molecular approaches such as DGGE, T-RFLP, 16S rRNA gene clone library analysis, and high-throughput sequencing methods (Baldi et al. 2010, Bowman et al. 2003, Bowman and McCuaig 2003, Carr et al. 2013, Carr et al. 2015, Fabiano and Danovaro 1998, Learman et al. 2016, Lee et al. 2014). Among the previous studies, only two studies applied high-throughput sequencing method of 16S rRNA genes for bacterial diversity analysis (Carr et al. 2013, Learman et al. 2016). However, the study by Carr et al (2013) was performed beneath the Ross Ice Shelf which persisted throughout the Holocene and thus, the sediments deposition was different with that from open area of the Ross Sea (Carr et al. 2013, D/'Hondt et al. 2015). In addition, Learman et al (2016) focused on the surface sediments remaining the bacterial diversity of subseafloor poorly understood.

In this context, bacterial diversity in the SO was investigated on the regional and local scale using pyrosequencing of the 16S rRNA gene. On the regional scale study,

bacterial communities across two latitudinal transects ranging from 47°S to 62°S and 50°S to 74°S in the SO across circumpolar fronts were investigated. Sampling with a higher spatial resolution around the PF and SAF boundary than previous studies to better understand for the main biogeographical boundaries structuring bacterial communities in the SO was performed. Furthermore, I examined whether the distribution of taxonomic groups was affected differently by the presence of each front and whether each frontal zone is characterized by certain species that characterize each frontal zone. On the local scale study, bacterial communities in seawater along the depth and distribution pattern of surface bacterial communities and spatial variability of bacterial community structure in marine sediments at many regions over the TNB of the Ross Sea were investigated. Environmental parameters explained best variability in the bacterial diversity were determined to understand key determinant that shape the bacterial communities.

2.2 Materials and Methods

Sampling information

For regional scale diversity study, surface water samples were collected at 25 stations in transect 1 ranging from 47°42.0307'S to 62°51.4661'S from February 27 to March 11, 2011 and at 17 stations in transect 2 ranging from 50°35.6532'S to 74°26.6614'S from January 18 to January 23, 2012 on the RV ARAON (Fig. 2-1). Sea surface temperature and salinity were measured while underway using a SBE45 thermosalinograph. The chlorophyll *a* concentration was measured using a continuous fluorescence sensor (10-AU, Turner Designs). At each station, 4 L of surface water was collected from the uncontaminated seawater intake system at a depth of about 7 m.

For local scale diversity study, seawater samples were collected at 20 stations in the TNB of the Ross Sea in February, 2011. At 10 stations (St.1–St.10), seawaters were collected using 24 Niskin Bottles (10 L each) equipped with CTD sensors (SBE 9; Sea-

Bird), and different sensor monitoring temperatures, salinities, dissolved oxygen concentrations, fluorescence levels, sampling depths, and pressures were assessed during the R/V ARAON expedition in the Ross Sea (Fig. 2-2). At 10 other stations (Z1–Z10), surface waters were collected using a 10-L bucket on the boat. Sampling depths at each station were as follows: 0, 10, 20, 30, 50, 75, 100, 150, and 200 m and then at 100-m intervals from 200 m to bottom (Table 2-1). Two liters of seawater were filtered through 3.0-μm cellulose acetate membranes (Advantec, Japan) and then the filtrate was filtered again through 0.2-μm polyethersulfone membranes (Pall, USA). The membranes were stored at –80°C for genomic DNA extraction. For nutrient analysis, untreated seawater was stored at –20°C until analysis. The positions of the each front were identified from a range of data recorded on board the R/V ARAON.

The shallow four sediment samples were collected using grab sampler on the boat in the TNB and the samples from deep water were collected using a dredge from two stations and gravity corer from one station on the R/V ARAON in the Ross Sea during the Austral summer of 2010-2012 and 2011-2012 (Fig. 2-3 and Table 2-2). The sampling depth ranged from 30.7 m to 1,224 m. For comparison, the biogenic surface sediments from two stations of Australian-Antarctic ridge with about 2,400 m of water depth were collected using a rock corer on the R/V ARAON in 2011 (Fig. 2-3 and Table 2-2). Portions of surface sediments (~1 cm) were collected aseptically from the box corer and grab sampler. Some portion of samples obtained using a dredge were taken and a whole round core about a 3.96-m long that was retrieved from the gravity corer was cut into half on board. Samples were taken with a sterile (autoclaved) spatula at 20 cm intervals including the surface. All samples were stored at -80°C until processing.

Chemical analysis

Concentrations of nitrite and nitrate, ammonium, phosphate, and silicate of seawater were analyzed using QuAAtro auto analyzer (SEAL Analytical, Germany) based on the manufacturer's instructions. Sediment samples obtained using a gravity core were analyzed for water content, grain size distribution, total nitrogen (TN) content and total inorganic carbon (TIC) content. The content of particles larger than sand-size was determined by wet sieving, and the content of silt- and clay-sized particles was analyzed using a Micrometric Sedigraph 5120. Contents of total nitrogen and total carbon (TC) were analyzed using a Flash EA 1112 element analyzer. Total organic carbon (TOC) content was determined by subtracting total inorganic carbon from total carbon content. Total inorganic carbon content was analyzed for the 2N HCl-dissolved portion using a UIC 5030 coulometer. The oxygen concentration in the sediment of this study was not measured. However, it was predictable that oxygen is not expected to be present below a few centimeters of sediments in the Ross Sea considering the high productivity in this region (Cai and Sayles 1996, D/'Hondt et al. 2015, Walker 2005) while surface sediment was considered to be oxic considering the oxygen level at the bottom depth of the water column in the Ross Sea (Fig. 2-4). Thus, the surface sample collected using a gravity core was considered to be oxic condition while the remainders were considered to be anoxic.

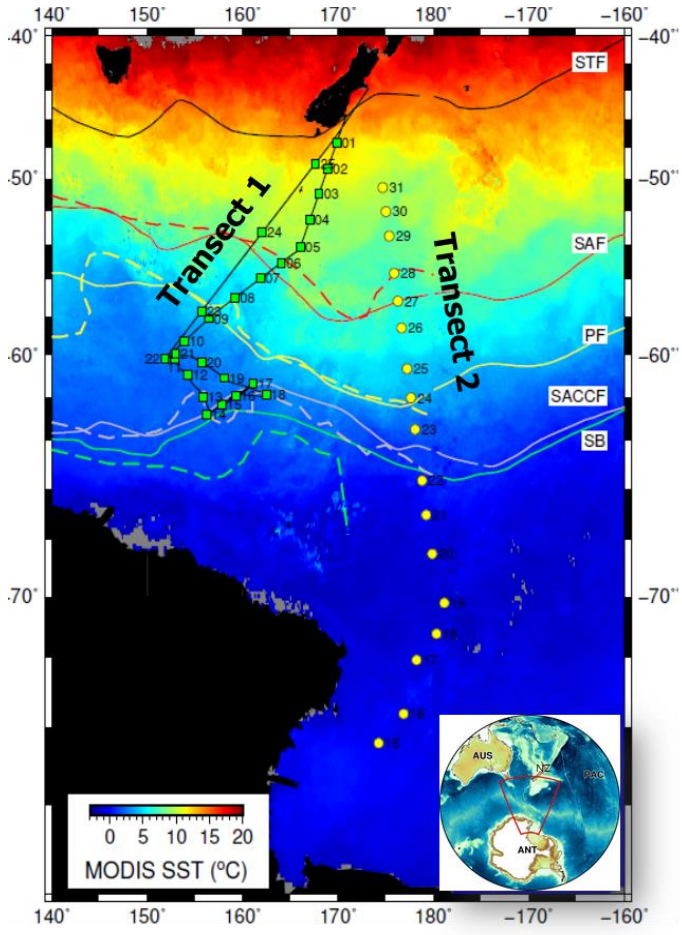


Figure 2-1. Map showing sampling locations superimposed on sea surface temperature of February, 2011. The SST was Level-3 mapped image derived from MODerate Resolution Imaging Spectroradiometer (MODIS). Approximate positions of the Subtropical Front (STF), the Subantarctic Front (SAF), the Polar Front (PF), the southern ACC front (SACCF), and the Southern Boundary (SB) are shown in colored lines. Southern Zone (SZ) lies between the SACCF and SB, and Subpolar Zone (SPZ) is the south of the SB. Inset: expanded view of the Southern Ocean including Australia (AUS), New Zealand (NZ), and Antarctica (ANT).

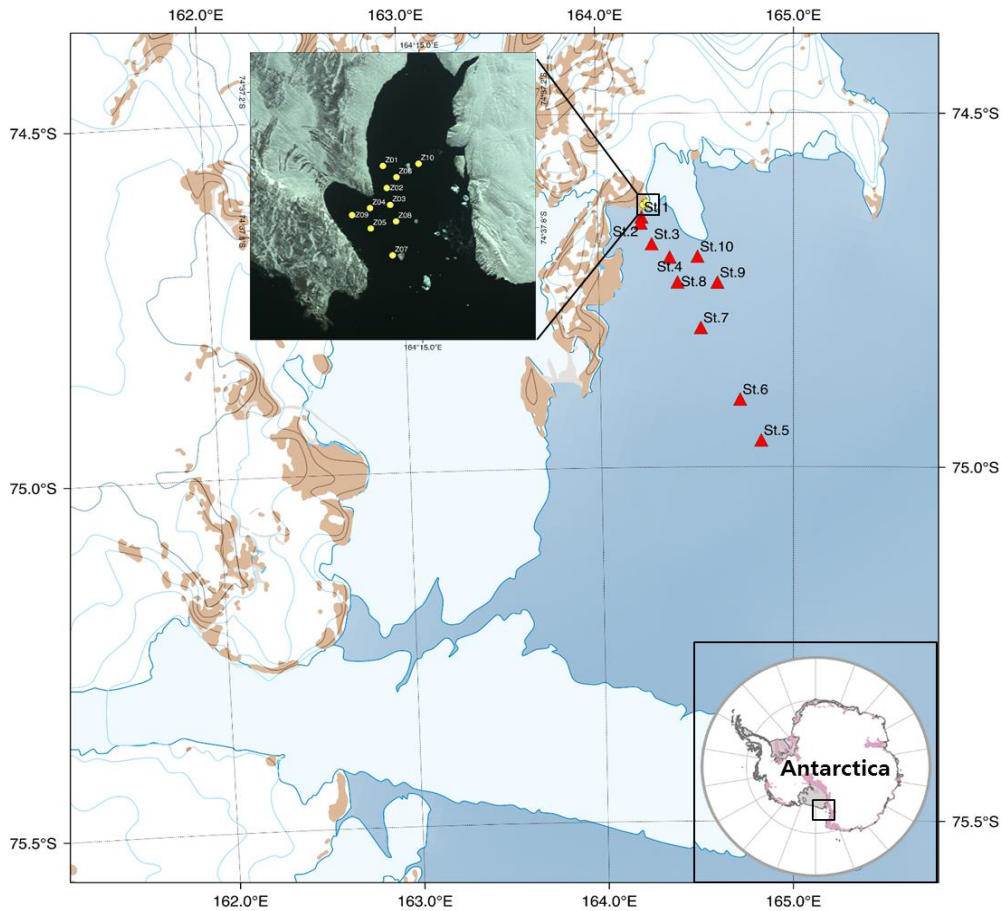


Figure 2-2. Sampling sites of seawater in the Terra Nova Bay, Ross Sea. Red triangles indicate stations that samples were collected using 24 Niskin Bottles on R/V ARAON and yellow circles indicate stations that samples were collected using a 10-L bucket on the boat. Inset: expanded view of Antarctica showing the Ross Sea.

Table 2-1. Location of sampling sites, water depth, sampling depth, environmental parameters and categorization of horizons

Station	Sampling depth (m)	Temperature (°C)	Salinity (psi)	Fluorescence	Oxygen	Pressure	PO_4^{3-} (mg/L)	$NO_3^- + NO_2^-$ (mg/L)	Si (mg/L)	NH_4^+ (mg/L)	Layer
St.1	0	0.25	33.57	4.34	404.48	1.54	0.018	0.025	0.987	0.018	Mixed layer
	10	0.26	33.59	4.59	400.73	10.16	0.021	0.048	1.172	0.015	Mixed layer
	20	0.25	33.65	1.14	363.13	20.31	0.035	0.178	2.329	0.017	Mixed layer
	30	-0.19	34.01	0.59	344.94	30.82	0.041	0.23	2.691	0.017	Thermocline
	50	-0.87	34.3	0.22	329.2	50.03	0.054	0.305	3.453	0.018	Thermocline
	75	-1.7	34.61	0.04	303.36	75.09	0.063	0.383	4.361	0.013	Deep water
	100	-1.8	34.63	0.07	293.62	100.38	0.064	0.4	4.673	0.013	Deep water
	150	-1.84	34.66	0.01	285.02	150.73	0.064	0.404	4.696	0.011	Deep water
St.2	0	0.37	33.59	4.46	386.53	1.73	0.019	0.03	1.018	0.014	Mixed layer
	10	0.31	33.66	1.93	391.83	10.24	0.024	0.078	1.475	0.014	Mixed layer
	20	0.15	33.79	1.02	357.53	19.95	0.038	0.192	2.318	0.044	Mixed layer
	30	-0.32	34.09	0.4	341.58	30.41	0.044	0.253	2.998	0.025	Thermocline
	50	-0.82	34.32	0.25	332.7	50.72	0.053	0.319	3.808	0.02	Thermocline
	75	-1.68	34.61	0.04	305.25	75.23	0.063	0.387	4.577	0.019	Deep water
	100	-1.79	34.63	0.04	292.2	100.37	0.066	0.409	4.647	0.015	Deep water
	150	-1.85	34.65	0.04	288	150.88	0.064	0.416	4.589	0.012	Deep water
	200	-1.88	34.67	0.04	285.04	200.32	0.066	0.433	4.776	0.012	Deep water

Table 2-1. To be continued

Station	Sampling depth (m)	Temperature (°C)	Salinity (psi)	Fluorescence	Oxygen	Pressure	PO_4^{3-} (mg/L)	$NO_3^- + NO_2^-$ (mg/L)	Si (mg/L)	NH_4^+ (mg/L)	Layer
St.3	0	0.16	33.66	4.1	378.75	1.47	0.018	0.099	1.21	0.02	Mixed layer
	10	0.12	33.68	4.31	375.09	10.75	0.019	0.106	1.243	0.014	Mixed layer
	20	0.07	33.75	4.04	368.66	20.41	0.02	0.117	1.294	0.012	Mixed layer
	30	-0.23	34.1	0.89	350.35	31.27	0.033	0.194	2.309	0.019	Thermocline
	50	-1.21	34.47	0.16	326.77	50.63	0.058	0.339	3.987	0.028	Thermocline
	75	-1.62	34.59	0.04	303.35	76.41	0.067	0.387	4.43	0.023	Deep water
	100	-1.75	34.62	0.04	295.36	102.3	0.063	0.408	4.655	0.02	Deep water
	150	-1.84	34.66	0.01	286.5	151.95	0.065	0.422	4.757	0.015	Deep water
	200	-1.88	34.67	0.04	284.72	200.82	0.065	0.426	4.827	0.013	Deep water
	300	-1.92	34.69	0.01	283.58	303.54	0.065	0.426	4.899	0.018	Deep water
	445	-1.95	34.71	0.01	284.58	450.84	0.064	0.425	4.874	0.015	Deep water
St.4	0	0.25	33.6	2.69	369.78	1.93	0.022	0.127	1.711	0.021	Mixed layer
	10	0.17	33.66	3.37	366.77	10.33	0.023	0.133	1.714	0.018	Mixed layer
	20	0	33.74	3.12	364.76	20.59	0.021	0.133	1.375	0.025	Mixed layer
	30	0.01	33.74	3.03	362.11	30.95	0.022	0.137	1.399	0.025	Mixed layer
	50	-0.89	34.28	0.53	321.27	51.04	0.055	0.333	3.493	0.031	Thermocline
	75	-1.63	34.59	0.04	307.03	76.05	0.064	0.392	3.996	0.021	Deep water
	100	-1.73	34.61	0.04	300.29	101.24	0.066	0.405	4.229	0.025	Deep water
	150	-1.85	34.64	0.04	285.62	151.67	0.068	0.429	4.705	0.021	Deep water
	200	-1.89	34.66	0.07	283.42	202.56	0.068	0.432	4.769	0.018	Deep water
	300	-1.9	34.68	0.04	283.77	303.5	0.065	0.433	4.474	0.015	Deep water
	335	-1.93	34.7	0.01	284.13	354.78	0.064	0.43	4.223	0.013	Deep water

Table 2-1. To be continued

Station	Sampling depth (m)	Temperature (°C)	Salinity (psi)	Fluorescence	Oxygen	Pressure	PO_4^{3-} (mg/L)	$NO_3^- + NO_2^-$ (mg/L)	Si (mg/L)	NH_4^+ (mg/L)	Layer
St.10	0	-1.28	33.25	1.23	366.76	1.99	0.017	0.138	1.394	0.01	Mixed layer
	10	-1.28	33.25	3.49	366.38	10.94	0.018	0.139	1.415	0.017	Mixed layer
	20	-1.05	33.41	3.64	357.98	20.22	0.022	0.164	1.643	0.018	Mixed layer
	30	-0.88	33.59	2.82	355.41	30.04	0.027	0.179	1.808	0.019	Mixed layer
	50	-0.48	33.91	1.47	345.65	50.9	0.035	0.207	2.16	0.02	Thermocline
	75	-1.43	34.55	0.16	311.49	75.78	0.059	0.35	3.971	0.013	Deep water
	100	-1.8	34.64	0.04	292.75	101.13	0.066	0.414	4.432	0.008	Deep water
	150	-1.81	34.66	0.04	288.84	151.77	0.069	0.423	4.577	0.006	Deep water
	200	-1.89	34.67	0.07	285.56	202.57	0.068	0.43	4.639	0	Deep water
	300	-1.93	34.7	0.01	284.08	303.92	0.068	0.433	4.643	0.003	Deep water
	420	-1.96	34.72	0.01	285.4	427.59	0.068	0.431	4.707	0	Deep water
St.8	0	-1.18	33.51	2.82	359.72	1.63	0.023	0.167	1.888	0.003	Mixed layer
	10	-1.14	33.52	3.61	358.79	10.19	0.024	0.168	1.886	-0.002	Mixed layer
	20	-1.09	33.55	3.7	357.08	20.65	0.026	0.173	1.922	0.004	Mixed layer
	30	-1.03	33.6	3.27	360.2	30.68	0.026	0.172	1.871	0.005	Mixed layer
	50	-0.21	33.72	2.08	356.05	51.03	0.025	0.148	1.516	0.01	Thermocline
	75	-1.34	34.51	0.16	323.18	75.81	0.057	0.337	3.789	0.016	Thermocline
	100	-1.52	34.56	0.07	311.57	100.38	0.06	0.381	4.165	0.01	Deep water
	150	-1.82	34.63	0.04	289.75	151.16	0.067	0.42	4.464	0.008	Deep water
	200	-1.87	34.65	0.04	284.85	202.22	0.064	0.43	4.526	0.003	Deep water
	300	-1.91	34.69	0.01	284.26	304.94	0.066	0.43	4.587	0.003	Deep water
	460	-1.99	34.72	0.01	287.88	465.96	0.065	0.425	4.716	0.003	Deep water

Table 2-1. To be continued

Station	Sampling depth (m)	Temperature (°C)	Salinity (psi)	Fluorescence	Oxygen	Pressure	PO_4^{3-} (mg/L)	$NO_3^- + NO_2^-$ (mg/L)	Si (mg/L)	NH_4^+ (mg/L)	Layer
St.9	0	-0.93	33.48	1.53	360.65	1.35	0.018	0.151	1.646	0.017	Mixed laver
	10	-0.93	33.48	3.03	360.36	10.15	0.022	0.151	1.642	0.018	Mixed layer
	20	-0.93	33.48	3.21	359.77	20.29	0.02	0.154	1.673	0.026	Mixed layer
	30	-0.92	33.48	3.09	360.51	29.89	0.02	0.155	1.679	0.028	Mixed layer
	50	-0.66	33.58	2.69	356.18	50.52	0.025	0.166	1.796	0.034	Thermocline
	75	-0.44	33.83	0.92	337.37	75.74	0.038	0.248	2.726	0.034	Thermocline
	100	-1.55	34.57	0.13	310.93	100.92	0.061	0.382	4.373	0.036	Deep water
	150	-1.81	34.63	0.04	290.66	151.61	0.067	0.427	4.732	0.026	Deep water
	200	-1.87	34.66	0.04	286.03	201.85	0.067	0.436	4.816	0.022	Deep water
	300	-1.94	34.7	0.01	285.51	303.83	0.067	0.439	4.869	0.021	Deep water
	400	-1.97	34.72	0.01	285.6	405.09	0.066	0.436	4.752	0.026	Deep water
	500	-1.94	34.73	0.01	284.52	505.36	0.051	0.306	2.812	0.052	Deep water
	590	-1.9	34.77	0.01	287.23	604.6	0.066	0.436	4.961	0.023	Deep water
St.7	0	-1	33.34	0.47	364.29	1.56	0.018	0.132	1.35	0.002	Mixed layer
	10	-0.99	33.34	2.27	364	10.83	0.019	0.134	1.363	0	Mixed layer
	20	-0.43	33.69	2.11	357.95	20.32	0.022	0.147	1.489	0.007	Mixed layer
	30	-0.42	33.7	2.36	357.63	31	0.025	0.165	1.683	0.009	Mixed layer
	50	-0.58	33.79	2.02	353.46	50.71	0.03	0.187	2.023	0.015	Thermocline
	75	-0.83	34.27	0.62	342.48	75.72	0.037	0.226	2.439	0.023	Thermocline
	100	-1.66	34.58	0.1	292.43	101.99	0.021	0.15	1.514	0.01	Deep water
	150	-1.86	34.63	0.04	285.37	152.19	0.065	0.425	4.617	0.01	Deep water
	200	-1.87	34.66	0.04	283.73	202.77	0.069	0.432	4.722	0.011	Deep water
	300	-1.89	34.7	0.01	282.04	303.55	0.069	0.432	4.722	0.011	Deep water
	400	-1.9	34.73	0.01	282.95	404.8	0.069	0.432	4.668	0.007	Deep water
	721	-1.89	34.8	0.01	283.34	506.1	0.064	0.425	4.871	0.002	Deep water

Table 2-1. To be continued

Station	Sampling depth (m)	Temperature (°C)	Salinity (psi)	Fluorescence	Oxygen	Pressure	PO_4^{3-} (mg/L)	$NO_3^- + NO_2^-$ (mg/L)	Si (mg/L)	NH_4^+ (mg/L)	Layer
St.6	0	-1.41	33.1	0.65	369.29	1.66	0.017	0.139	1.6	0.008	Mixed layer
	10	-1.38	33.12	3.79	368.54	10.45	0.019	0.15	1.661	0.01	Mixed layer
	20	-1.03	34.29	0.86	332.18	19.95	0.047	0.278	2.968	0.027	Thermocline
	30	-1.41	34.45	0.44	313.18	30.92	0.057	0.355	3.778	0.025	Thermocline
	50	-1.67	34.56	0.16	294.15	50.99	0.062	0.403	4.322	0.021	Deep water
	75	-1.77	34.6	0.1	288.87	76.39	0.062	0.412	4.383	0.021	Deep water
	100	-1.8	34.62	0.07	290.45	101.96	0.062	0.42	4.017	0.018	Deep water
	150	-1.86	34.64	0.07	283.22	151.67	0.069	0.435	4.476	0.02	Deep water
	200	-1.87	34.65	0.04	282.7	202.66	0.067	0.434	4.649	0.019	Deep water
	300	-1.88	34.69	0.04	280.13	303.88	0.066	0.436	4.616	0.022	Deep water
	400	-1.89	34.72	0.01	280.22	405.32	0.065	0.434	4.599	0.022	Deep water
	500	-1.89	34.74	0.01	282.32	505.91	0.066	0.433	4.615	0.02	Deep water
	640	-1.9	34.77	0.01	287.06	648.4	0.065	0.426	4.526	0.02	Deep water
St.5	0	-1.06	33.59	1.17	354.69	1.62	0.026	0.184	1.948	0.025	Mixed layer
	10	-0.95	34.45	0.86	317.31	6.49	0.058	0.338	3.822	0.029	Mixed layer
	20	-1.52	34.54	0.25	306.68	19.57	0.063	0.383	3.965	0.033	Thermocline
	30	-1.71	34.59	0.16	300.62	29.12	0.06	0.396	4.338	0.028	Thermocline
	50	-1.77	34.61	0.07	291.46	51.04	0.06	0.396	4.338	0.028	Deep water
	75	-1.86	34.64	0.07	285.79	70.66	0.066	0.428	4.681	0.023	Deep water
	100	-1.9	34.65	0.04	284.39	101.33	0.067	0.433	4.554	0.022	Deep water
	150	-1.89	34.67	0.04	282.39	151.63	0.067	0.433	4.554	0.022	Deep water
	200	-1.89	34.69	0.07	281.85	201.86	0.067	0.434	4.53	0.02	Deep water
	916	-1.88	34.8	0.01	287.8	928.06	0.066	0.419	4.804	0.015	Deep water

Table 2-1. To be continued

Station	Sampling depth (m)	Temperature (°C)	Salinity (psi)	Fluorescence	Oxygen	Pressure	PO_4^{3-} (mg/L)	$NO_3^- + NO_2^-$ (mg/L)	Si (mg/L)	NH_4^+ (mg/L)	Layer
Z01		-0.2	15	nd*	nd	nd	0.02	0.087	1.31	0.025	Mixed layer
Z02		0.1	34	nd	nd	nd	0.015	0.063	1.088	0.023	Mixed layer
Z04		0	34.1	nd	nd	nd	0.021	0.062	1.091	0.025	Mixed layer
Z03		0	33.4	nd	nd	nd	0.02	0.076	1.209	0.026	Mixed layer
Z05		-0.1	33.8	nd	nd	nd	0.016	0.068	1.148	0.023	Mixed layer
Z06		-0.2	34	nd	nd	nd	0.017	0.074	1.175	0.025	Mixed layer
Z08		-0.2	33.3	nd	nd	nd	0.017	0.062	1.059	0.03	Mixed layer
Z09		0	33.9	nd	nd	nd	0.016	0.063	1.052	0.025	Mixed layer
Z10		-0.4	33.2	nd	nd	nd	0.018	0.074	1.175	0.027	Mixed layer
Z07		-0.4	34.1	nd	nd	nd	0.018	0.069	1.054	0.029	Mixed layer

* nd indicates not measured.

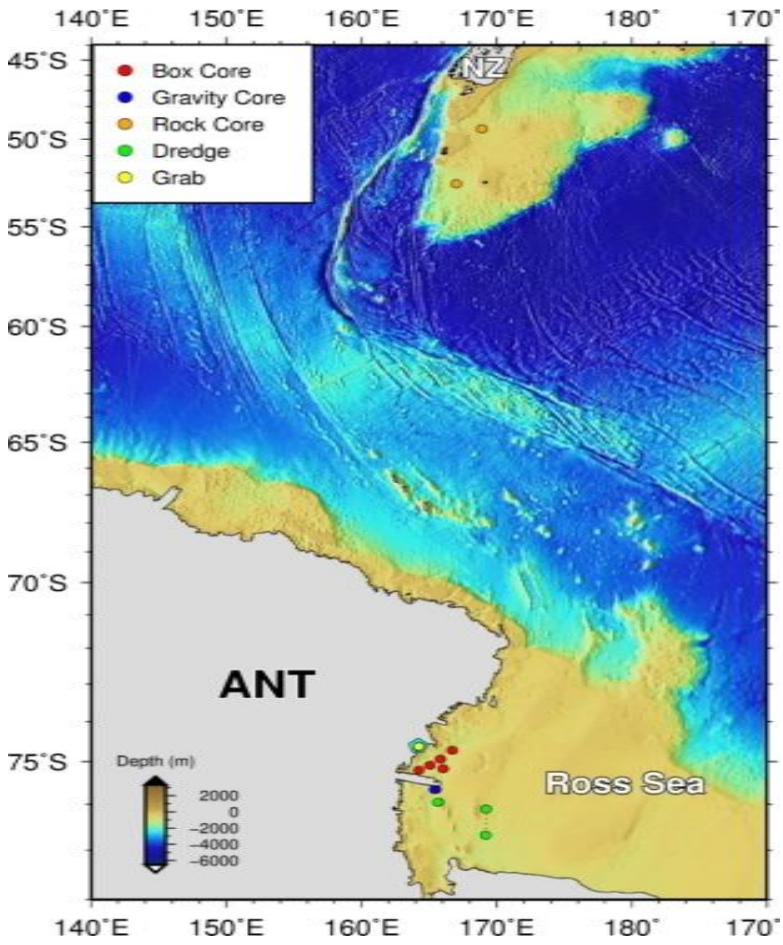


Figure 2-3. Sampling sites of marine sediments

Table 2-2. Sapling sites of marine sediments information

Location	Sample no.	Latitude	Longitude	Depth (m)	Sampling method
Ross Sea	LYM110206-01	74°38'46.20"S	164°13'24.60"E	-	Box corer
	LYM110208-01	76°06'17.10"S/ 169°12'45.36"E - 76°41'03.60"S/ 169°11'30.66"E	-	-	Dredge
	LYM110211-01	74.6251695" S	164.2487574" E	54	Grab sampler
	LYM110211-05	74.6261446" S	164.2433630" E	-	Grab sampler
	LYM110211-12	74.6298891" S	164.2366488" E	30.7	Grab sampler
	LYM110211-14	74.6316075" S	164.2416185" E	43.1	Grab sampler
	DG12-GC06	75°39.5684' S	165° 23.8382' E	859	Gravity core
	NHJ120101-01	75° 12.6821' S	164 14.1323' E	1224	Box corer
	NHJ120102-01	74° 56.3592' S	165 50.4515' E	1128	Box corer
	NHJ120109-01	74° 43.6415' S	166 42.2172' E	1028	Box corer
	NHJ120113-01	75° 10.1984' S	166° 2.1842' E	825	Box corer
	NHJ120113-02	75° 5.3106' S	165° 3.4404' E	1165	Box corer
	NHJ120114-01	75° 56' 59.465" S/ 165 41' 30.318 E - 75° 57' 16.908" S, 165 39' 08.700" E	-	517	Dredge
	NHJ120114-02	75° 39.5684‘ S	165° 28.8382' E	859	Box corer
AAR*	KRR1-RC12-S	49° 22.8994' S	168° 54.1902' E	2,479	Rock corer
	KRR1-RC14-S	52° 37.1942' S	167° 00.7434' E	2,359	Rock corer

* AAR: Australian-Antarctic Ridge

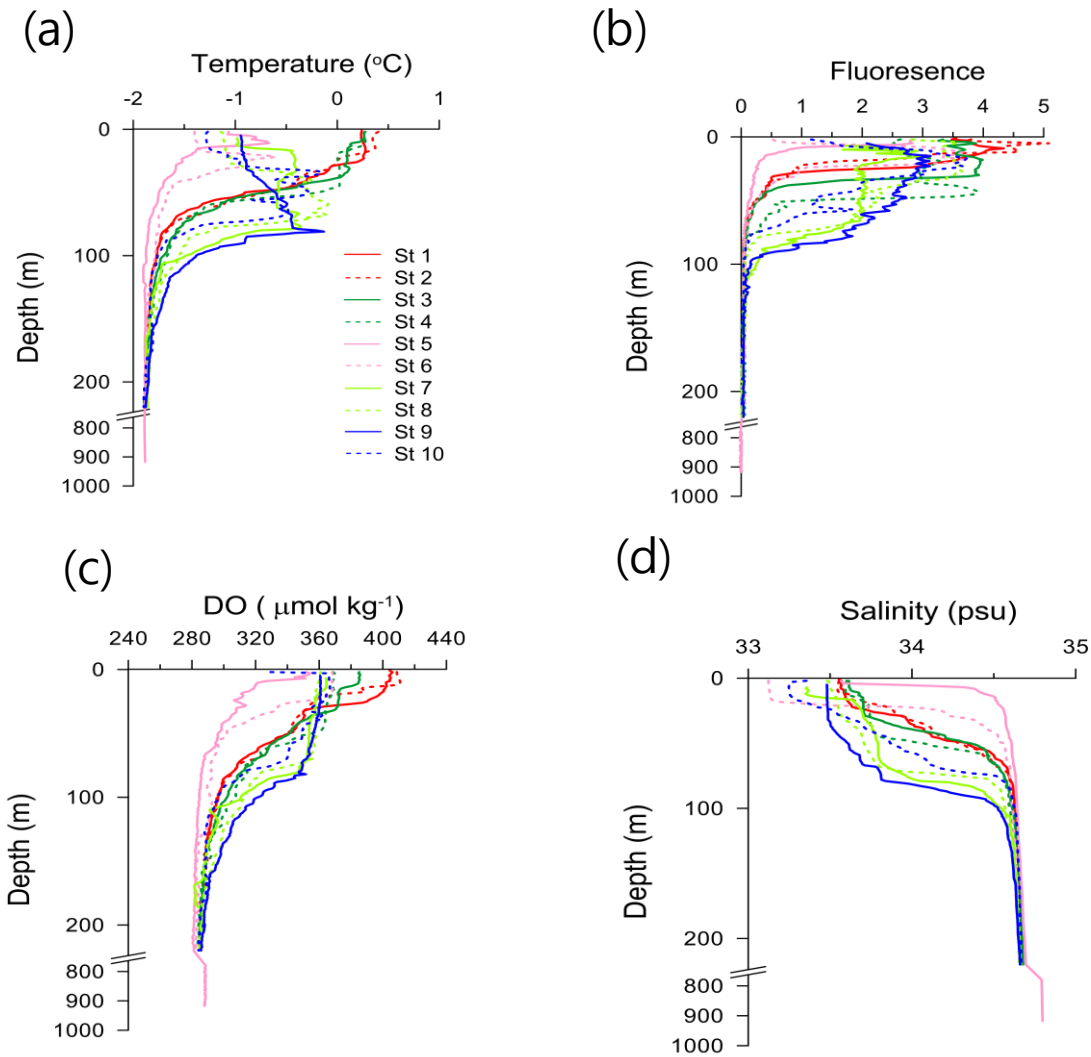


Figure 2-4. Vertical profiles of temperature (a), fluorescence (b), dissolved oxygen (c), and salinity (d).

Genomic DNA extraction and PCR amplification

Total genomic DNA was extracted from the 0.2-μm membranes or approximately 0.5 g of each sediment sample using FastDNA SPIN Kit for Soil (Q-Biogene, USA) according to the manufacturer's instructions. DNA concentrations were determined using a Nanodrop 8000 Spectrophotometer (Thermo Scientific, USA). Bacterial 16S rRNA gene sequences were amplified with 27F (Lane 1991) and 518R (Kato et al. 1997) barcode primers. PCR was performed in 30 μl reaction mixtures containing 10 X PCR reaction buffer, 160 μM dNTPs, 0.5 μM each primer, approximately 10 ng gDNA, and 2.5 unit *Taq* DNA polymerase (GeneAll, Korea). The PCR procedure for seawater samples included an initial denaturing step at 94°C for 3 min, 25 cycles of amplification (94°C for 1 min, 55°C for 1 min, and 72°C for 1.5 min), and final extension step at 72°C for 7 min and for sediment sample an initial denaturation step at 94°C for 3 min, 30 cycles of amplification (94°C for 1 min, 55°C for 1 min, and 72°C for 1 min 30s), and a final extension step at 72°C for 5 min. Each sample was amplified in triplicate and pooled to reduce PCR bias. PCR products were purified using the LaboPass Purification Kit (Cosmogentech, Korea).

Pyrosequencing and sequence analysis

Sequencing of the 16S rDNA amplicon was carried out by DNA Link and Chun Lab (Korea) using a 454 GS FLX-Titanium sequencing instrument (Roche, USA). Pre-processing was conducted using PyroTrimmer (Oh et al. 2012). Sequences were processed to remove primer, linker, and barcode sequences. The 3′ ends of sequences with low quality values were trimmed when average quality scores for a 5-bp window size were lower than 20. Sequences with ambiguous nucleotides or that were <200 bp were discarded. Sequence clustering was performed with CLUSTOM (Hwang et al. 2013) using a 97% similarity cutoff. Chimeric reads were detected and discarded using the *de novo* chimera detection algorithm of UCHIME (Edgar et al. 2011). Taxonomic assignment was conducted for

representative sequences of each cluster by EzTaxon-e database (Kim et al. 2012b)

Diversity estimation and statistical analysis

Relative differences among samples due to environmental variables such as salinity, temperature, fluorescence, and concentration of nitrite and nitrate, ammonium, phosphate, and silicate were calculated by determining Euclidean distances using PRIMER software, version 6 (Clarke and Gorley 2006). Diversity indices, including operational taxonomic unit (OTU) richness, Chao1, and ACE indices were estimated using Mothur (Schloss et al. 2009) based on OTUs, sequence groups with more than 97% sequence similarity.

Bray-Curtis similarities were calculated using a Hellinger-transformed OTU abundance matrix and the relative similarities of bacterial communities among samples were visualized by non-metric multidimensional scaling (NMDS). To determine the relationship between environmental variables and bacterial community structure, vectors of environmental variables were fitted onto the ordination space the "ordisurf" and "envfit" functions in the vegan R package, and the significance of each correlation was tested based on 999 permutations (Oksanen et al. 2013). Prior to fitting vectors onto the ordination space, multicollinearity between variables was tested using variance inflation factors and skewed variables were transformed based on 'Draftsman plot' result in PRIMER v6 (Clarke and Gorley 2006). For example, NH_4^+ concentration was log-transformed, fluorescence was square root transformed, and other variables were left unchanged. An analysis of similarity (ANOSIM) was performed with 999 permutations to test if bacterial community structures are significantly differentiated by each frontal zone. The relative abundance of each phylum and major OTUs (≥2% in relative abundance) were correlated to latitudinal change using Pearson's product-moment correlation to see if there are any positive or negative relationships between them.

2.3 Results

2.3.1 Biogeographic structure of the bacterial communities according to the circumpolar fronts in the Southern Ocean

Change of environmental factors and correlation among environmental factors

The temperature decreased from 12.9°C to 2.8°C throughout transect 1 and from 10.3°C to 3.0°C throughout the transect 2 as the latitude increased (Fig. 2-5 and Table 2-3). The salinity ranged from 33.7 to 34.4 and decreased with increasing latitude (Fig. 2-5 and Table 2-3). Nitrite, nitrate, phosphate, and silicate concentrations increased as the latitude increased (Fig. 2-5 and Table 2-3). Similarity analysis by Euclidean distance based on nutrient concentrations, temperature, salinity, and fluorescence showed that samples were divided by the transect (except for D-24) and frontal systems along the same transect (except for T-24), as shown in Fig. 2-6. Samples in the SAZ and SPZ were especially distinguished from those of other zones in each transect, based on environmental factors (Fig. 2-6).

Analysis of environmental variables such as latitude; phosphate, nitrite, nitrate, silicate, and ammonium concentrations; temperature; salinity; and fluorescence revealed strong correlations among them: the latitude was significantly positively correlated with phosphate, nitrite, nitrate, and silicate concentrations, and the latitude was negatively correlated with temperature and salinity (Table 2-4).

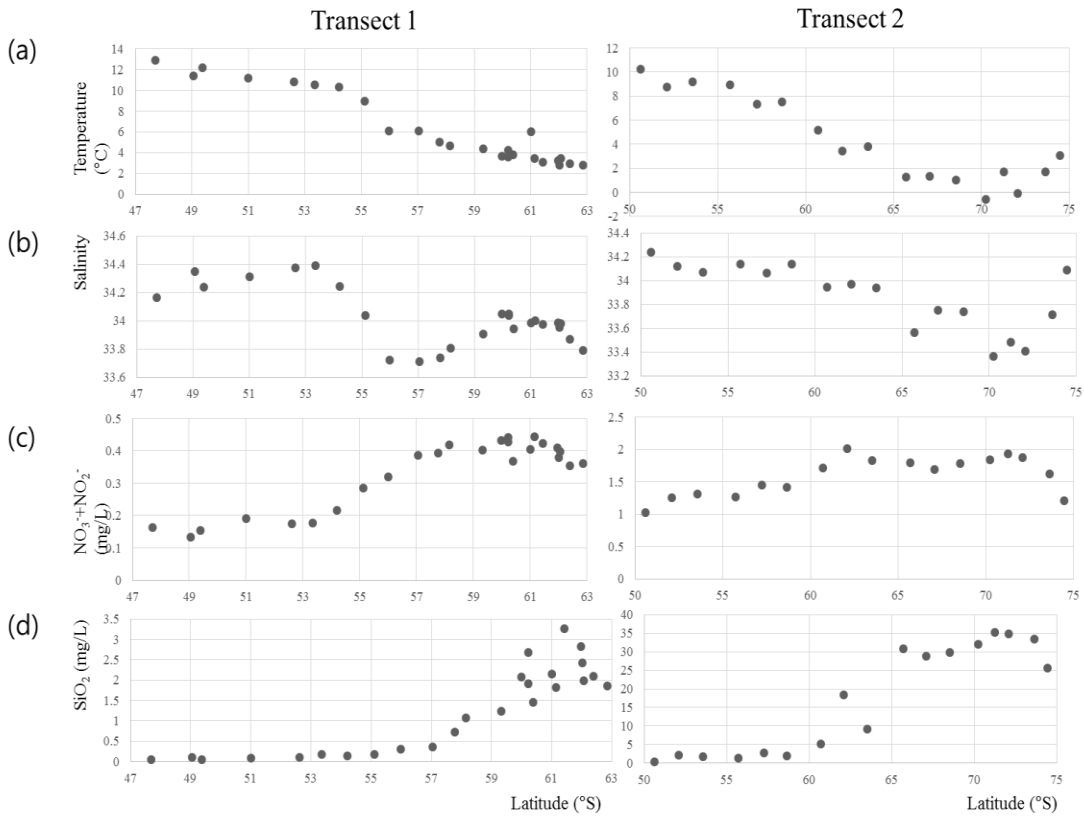


Figure 2-5. Temperature (a), salinity (b), and concentration of nitrite and nitrate (c) and silicate (d) profiles along the latitude in transect 1 and transect 2.

Table 2-3. Location of sampling sites and environmental variables

Transect	Station ID*	Latitude	Longitude	Frontal zone[§]	PO_4 (mg/L)	NO_3+NO_2 (mg/L)	SiO_2 (mg/L)	NH_4 (mg/L)	Temperature (°C)	Salinity (psu)	Fluorescence
	D-01	47° 42.0307' S	169° 49.9920' E	SAZ	0.028	0.164	0.043	0.003	12.9211	34.1646	0.839
	D-25	49° 03.0677' S	167° 36.7994' E	SAZ	0.021	0.134	0.099	0	11.3835	34.348	0.654
	D-02	49° 22.8994' S	168° 54.1902' E	SAZ	0.031	0.155	0.039	0.003	12.1778	34.2394	0.847
	D-03	50° 59.8212' S	167° 58.6344' E	SAZ	0.033	0.191	0.079	0.003	11.1664	34.3125	0.319
	D-04	52° 37.1942' S	167° 00.7434' E	SAZ	0.029	0.175	0.1	0.003	10.8232	34.3748	0.956
	D-24	53° 20.9974' S	162° 02.7585' E	SAZ	0.032	0.176	0.167	0.001	10.5323	34.3903	2.693
	D-05	54° 12.0708' S	166° 02.1362' E	SAZ	0.04	0.217	0.133	0.004	10.2926	34.2413	0.761
	D-06	55° 06.7618' S	164° 02.3146' E	SAZ	0.045	0.284	0.172	0.003	8.9628	34.0359	0.471
	D-07	55° 59.3785' S	161° 52.3106' E	PFZ	0.05	0.32	0.3	0.002	6.1113	33.7209	0.566
	D-08	57° 02.7004' S	159° 12.1593' E	PFZ	0.057	0.386	0.353	0.002	6.066	33.7096	0.711
	D-23	57° 46.0923' S	155° 41.7314' E	AZ	0.062	0.394	0.72	0.004	5.0432	33.7365	0.651
1	D-09	58° 09.0192' S	156° 27.5763' E	AZ	0.064	0.418	1.07	0	4.6508	33.8062	0.411
	D-10	59° 18.6712' S	153° 55.1596' E	AZ	0.057	0.401	1.243	0.001	4.3851	33.9033	0.745
	D-21	59° 58.9188' S	152° 56.8440' E	AZ	0.055	0.433	2.077	0.003	3.6644	34.0463	0.645
	D-22	60° 12.6852' S	151° 50.9143' E	AZ	0.076	0.441	2.679	0.061	3.5736	34.0486	0.489
	D-11	60° 12.8018' S	152° 51.6966' E	AZ	0.069	0.427	1.921	0.06	4.1962	34.0364	0.464
	D-20	60° 23.0720' S	155° 41.9033' E	AZ	0.058	0.368	1.462	0.004	3.8296	33.9415	0.575
	D-12	61° 00.4621' S	154° 17.0690' E	AZ	0.055	0.404	2.15	0.003	6.0056	33.9844	1.254
	D-19	61° 08.8424' S	158° 04.9851' E	AZ	0.07	0.443	1.816	0.008	3.4621	33.998	0.568
	D-17	61° 25.7009' S	161° 04.8402' E	AZ	0.067	0.423	3.267	0.033	3.1014	33.9755	2.164
	D-13	62° 03.6210' S	155° 50.7834' E	AZ	0.058	0.397	1.99	0.004	3.4182	33.9797	1.16
	D-18	61° 57.7393' S	162° 26.8809' E	SZ	0.054	0.41	2.824	0	3.2347	33.9847	0.877
	D-16	62° 00.4497' S	159° 17.5959' E	SZ	0.055	0.378	2.424	0.002	2.814	33.9506	1.172

Table 2-3. To be continued

Transect	Station ID*	Latitude	Longitude	Frontal zone§	PO_4 (mg/L)	NO_3+N (mg/L)	SiO_2 (mg/L)	NH_4 (mg/L)	Temperature (°C)	Salinity (psu)	Fluorescence
1	D-15	62° 23.2354'S	157° 49.5763' E	SZ	0.053	0.354	2.1	0.005	2.9628	33.8693	4.057
	D-14	62° 51.4661' S	156° 12.4310' E	SZ	0.055	0.36	1.849	0.007	2.8148	33.7885	1.806
2	T-31	50° 35.6532' S	174° 39.5775' E	SAZ	0.145	1.028	0.277	0.125	10.2654	34.2366	1.381
	T-30	52° 05.5341' S	174° 59.8776' E	SAZ	0.172	1.253	2.132	0.125	8.7601	34.1219	0.924
	T-29	53° 33.8079' S	175° 20.6020' E	SAZ	0.177	1.31	1.721	0.12	9.153	34.0668	1.038
	T-28	55° 42.4323' S	175° 51.8592' E	SAZ	0.171	1.261	1.443	0.144	8.9074	34.1361	1.855
	T-27	57° 14.1478' S	176° 15.4663' E	SAZ	0.196	1.454	2.743	0.162	7.3467	34.066	1.903
	T-26	58° 38.9255' S	176° 38.2365' E	PFZ	0.188	1.411	2.044	0.222	7.5291	34.1372	2.594
	T-25	60° 42.1046' S	177° 12.6321' E	PFZ	0.232	1.712	5.127	0.12	5.146	33.9447	4.057
	T-24	62° 05.1296' S	177° 37.1609' E	PFZ	0.259	2.008	18.447	0.187	3.41	33.9716	3.884
	T-23	63° 31.1609' S	178° 03.6607' E	AZ	0.24	1.825	9.222	0.124	3.7975	33.9376	4.42
	T-22	65° 41.6534' S	178° 46.9381' E	SPZ	0.234	1.799	30.882	0.127	1.2463	33.5605	3.683
	T-21	67° 03.9282' S	179° 13.3155' E	SPZ	0.221	1.694	28.805	0.113	1.3209	33.7477	3.801
	T-20	68° 31.5192' S	179° 51.2626' E	SPZ	0.241	1.78	29.886	0.104	1.0401	33.7352	4.438
	T-19	70° 14.6886' S	178° 53.1755' E	SPZ	0.247	1.842	32.056	0.104	-0.5677	33.3635	4.695
	T-18	71° 15.3544' S	179° 41.9595' E	SPZ	0.258	1.936	35.182	0.109	1.7065	33.4824	3.614
	T-17	72° 04.3347' S	178° 12.9320' E	SPZ	0.254	1.87	34.779	0.093	-0.1066	33.4063	3.38
	T-16	73° 38.6780' S	176° 51.2060' E	SPZ	0.226	1.622	33.463	0.123	1.7198	33.715	2.555
	T-15	74° 26.6614' S	174° 13.7154' E	SPZ	0.193	1.205	25.632	0.082	3.0571	34.0883	2.07

* Station ID was designated with '-D' for stations along transect 1 and '-T' for stations along transect 2 of Figure 2-1.

§Acronyms: SAZ: Subantarctic Zone lies between the Subtropical Front and Subantarctic Front, PFZ: Polar Frontal Zone lies between the Subantarctic Front and the Polar Front, AZ: Antarctic Zone lies between the Polar Front and the southern ACC front, SZ: Southern Zone lies between the southern ACC front and Southern Boundary, and SPZ: Subpolar Zone, the south of the SB.

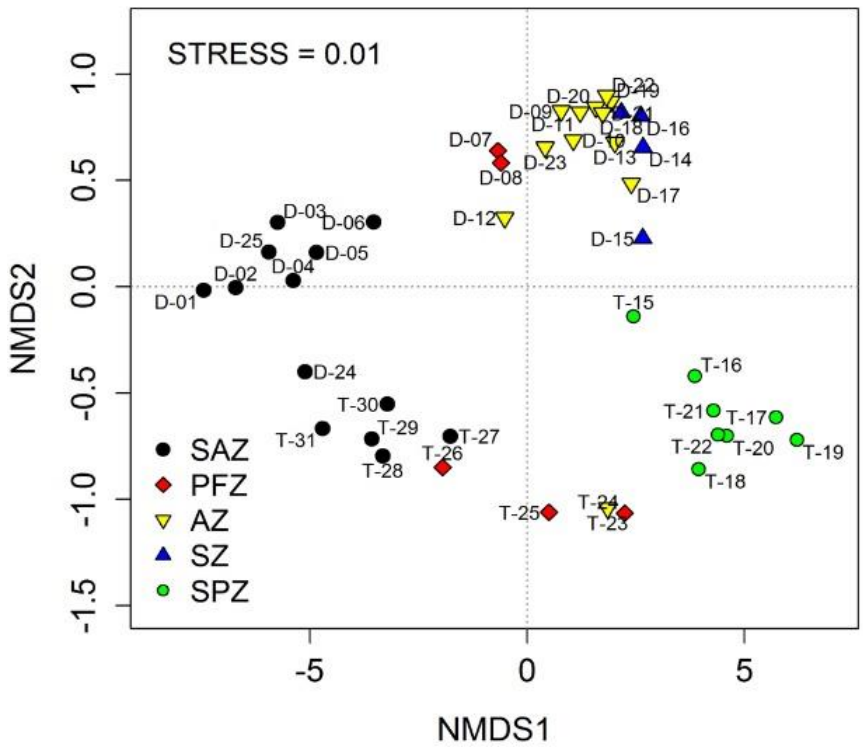


Figure 2-6. Similarities of samples by environmental variables. The Euclidean distance was calculated using the values of environmental variables, salinity, temperature, fluorescence, and concentration of nitrite and nitrate, ammonium, phosphate, and silicate.

Bacterial diversity along the fronts

Of the twenty eight bacterial phyla or classes of *Proteobacteria* recovered, *Bacteroidetes*, *Alphaproteobacteria*, *Gammaproteobacteria*, *Betaproteobacteria*, *Cyanobacteria*, *Verrucomicrobia*, and SAR406 were predominant composing more than 97.2% of the bacterial community across all stations (Fig. 2-7a and b). In particular, one phylum (*Bacteroidetes*) and two proteobacterial classes (*Alphaproteobacteria* and *Gammaproteobacteria*) dominated almost all samples, consisting of 82.5-99.8% of the total bacterial communities. Significant differences in the levels of diversity were found between samples when comparing diversity indices, such as the Chao1, ACE, Shannon, and Simpson indices (Table 2-7). Correlation analysis between the diversity indices and latitude revealed that bacterial diversity (represented by OTU numbers and the Chao1, ACE, and Shannon diversity indices) decreased as the latitude increased (Table 2-8).

To investigate if the distribution of each bacterial phylum was influenced differently by the presence of the fronts, the correlation between the relative abundance of each phylum and latitude was analyzed by calculating Pearson's correlation coefficients versus the location of each front. In transect 1, the relative abundance of *Bacteroidetes* showed a significant positive correlation with the latitude, whereas *Cyanobacteria*, *Actinobacteria*, *Verrucomicrobia*, and SAR406 displayed the opposite trend (Table 2-5 and Fig. 2-8). *Bacteroidetes*, which were abundant across all stations with a range of 25.3–56.9% in relative abundance, showed significant and positive Pearson's correlation coefficients ($r = 0.76$, $P < 0.001$). Regardless of the increase of *Bacteroidetes* with increasing latitude, it was not clear whether any frontal system significantly affected the distribution of *Bacteroidetes* (Table 2-6, Fig. 2-7a and 2-8a).

Table 2-4. Pearson correlation coefficients among environmental variables

Environmental	Latitude		PO_4^{3-}		$NO_3^- + NO_2^-$		SiO_2		NH_4^+		Temperature		Salinity		Fluorescence	
factor	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2
Latitude	-															
PO_4^{3-}	0.85***	0.72***	-													
$NO_3^- + NO_2^-$	0.90***	0.59*	0.96***	0.98***	-											
SiO_2	0.86***	0.92***	0.74***	0.74***	0.79***	0.65*	-									
NH_4^+	0.27	-0.47	0.53**	-0.18	0.36	-0.06	0.45*	-0.51*	-							
Temperature	-	-0.91***	-0.89***	-0.87***	-0.94***	-0.79***	-0.85***	-0.94***	-0.29	0.44	-					
Salinity	-0.65***	-0.74***	-0.70***	-0.80***	-0.74***	-0.77***	-0.39	-0.87***	-0.01	0.45	0.73***	0.88***	-			
Fluorescence	0.26	0.60*	-0.05	0.87***	-0.03	0.87***	0.27	0.59*	-0.07	-0.14	-0.17	-0.79***	-0.01	-0.67**	-	

§ T1 and T2 represent transect 1 and transect 2, respectively.

* indicates p-value <0.05, ** p-value <0.01, and *** p-value <0.001, respectively.

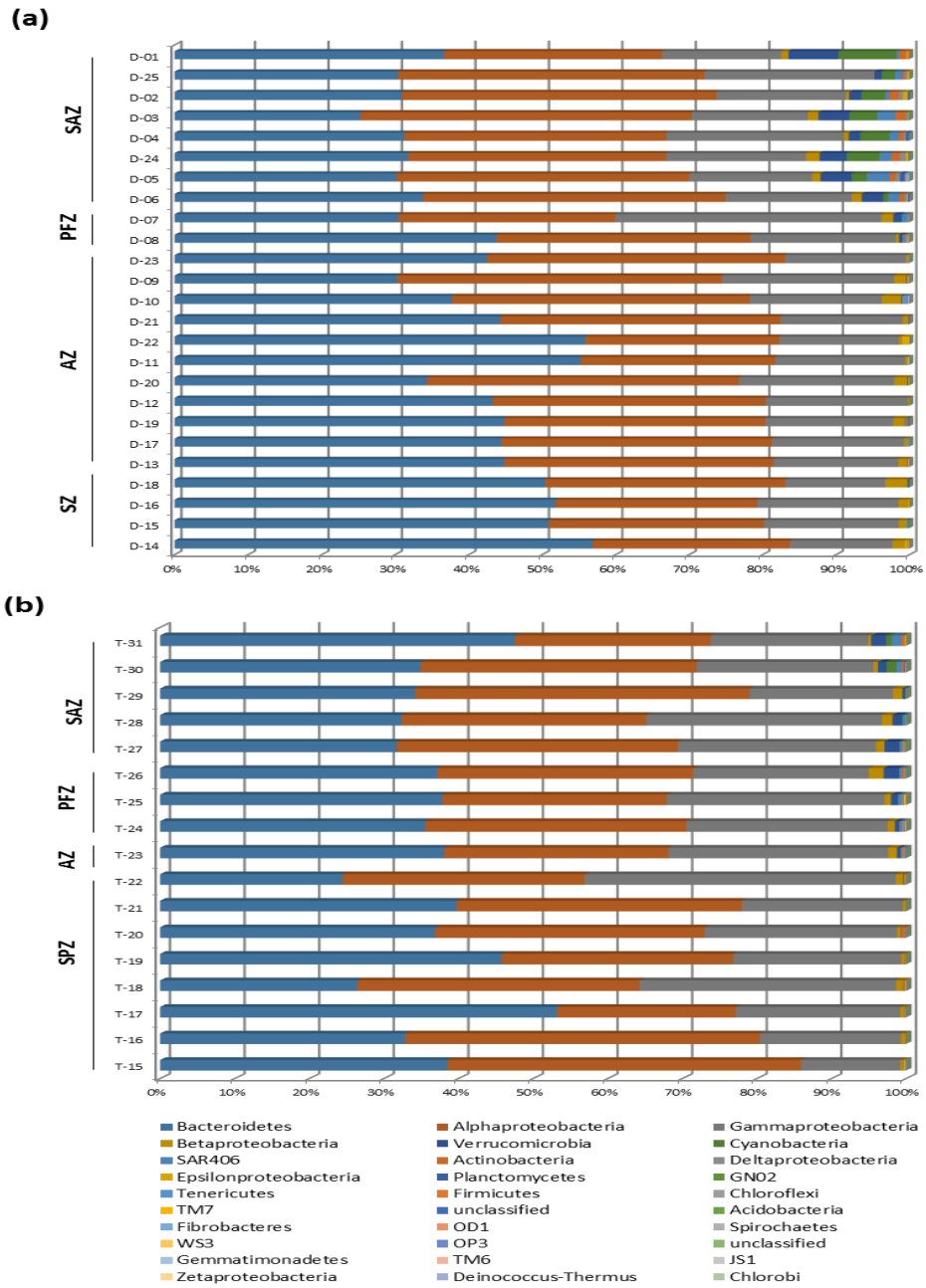


Figure 2-7. Relative abundance of bacterial communities at the phylum or class level of *Proteobacteria* in transect 1 (a) and transect 2 (b).

Table 2-5. Pearson correlation coefficients between latitudinal change and relative abundance of each bacterial phylum and class of *Proteobacteria*

Phylum or class of *Proteobacteria*	r1[†]	r2[§]
Bacteroidetes	**0.76*** **	0.38
GN02	0.34	0.37
Deinococcus-Thermus	0.29	0.29
OD1	0.24	0.11
Betaproteobacteria	0.21	-0.56
WS3	-0.02	-0.02
Chloroflexi	-0.15	0.29
Fibrobacteres	-0.15	-0.37
Gammaproteobacteria	-0.19	-0.28
Epsilonproteobacteria	-0.19	-0.06
OP3	-0.22	-0.32
Acidobacteria	-0.29	-0.22
Planctomycetes	-0.31	-0.29
Alphaproteobacteria	-0.38	-0.09
Firmicutes	-0.44*	-0.13
Deltaproteobacteria	-0.57**	-0.2
SAR406	**-0.62****	**-0.75*****
Cyanobacteria	**-0.78*****	-0.59**
Verrucomicrobia	**-0.78*****	**-0.79*****
Chlorobi	-	0.3
JS1	-	-0.3
TM7	-	-0.35
Zetaproteobacteria	-	-0.07
Actinobacteria	**-0.80*****	-0.23
Gemmatimonadetes	0.21	-
Spirochaetes	0.12	-
Tenericutes	0.08	-
TM6	0.08	-

† Pearson correlation coefficient in transect 1 and § Pearson correlation coefficient in transect 2

* p-value <0.05, ** p-value <0.01, and *** p-value <0.001

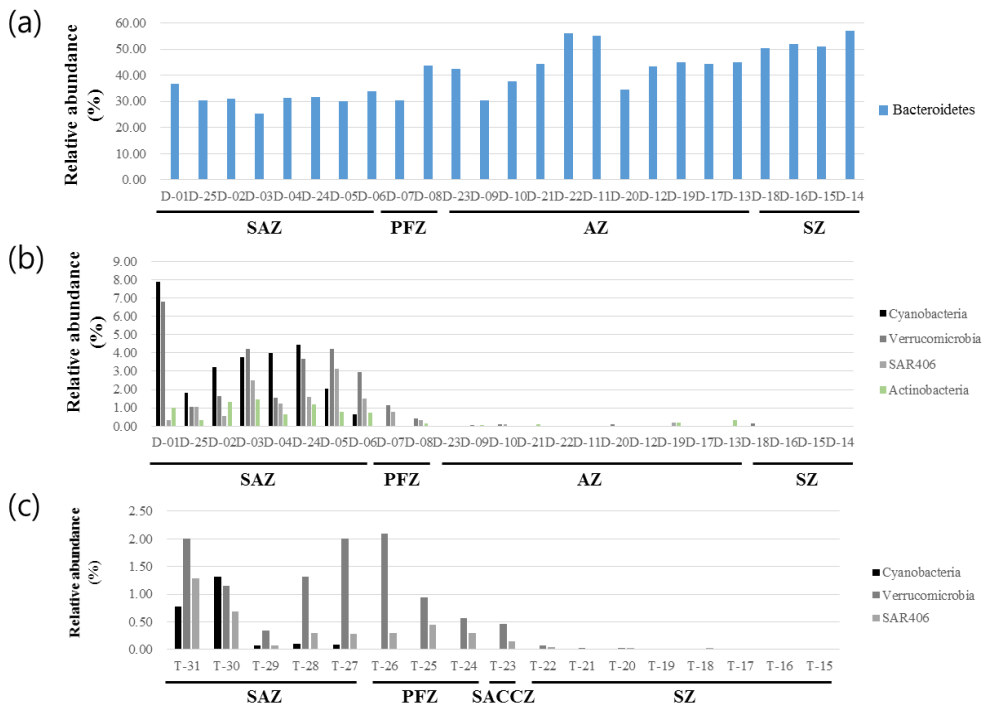


Figure 2-8. Phyla showing significantly different proportion along the front in transect 1 (a) and (b) and transect 2 (c).

Table 2-6. Relative abundance of each bacterial phylum or class of *Proteobacteria* at each frontal zone

Phylum or class of Proteobacteria	Relative abundance							
	Transect 1				Transect 2			
	SAZ	PFZ	AZ	SZ	SAZ	PFZ	AZ	SPZ
Bacteroidetes	31.26	37.11	43.43	52.5	36.14	36.85	38.05	37.24
Alphaproteobacteria	38.9	32.03	36.79	29.07	35.73	33.14	30.11	36.81
Gammaproteobacteria	18.71	28.01	18.33	16.29	24.44	26.58	29.45	25.06
Betaproteobacteria	1.01	1.06	1	1.76	0.97	1.27	1.18	0.57
Verrucomicrobia	3.27	0.78	0.02	0.04	1.36	1.2	0.47	0.02
Cyanobacteria	3.48	0	0	0	0.47	0	0	0
SAR406	1.49	0.55	0.03	0	0.52	0.35	0.15	0.01
Actinobacteria	0.94	0.08	0.06	0	0.17	0.12	0.2	0.16
Deltaproteobacteria	0.42	0.27	0.1	0.04	0.05	0.33	0.34	0.04
Epsilonproteobacteria	0.2	0.04	0.11	0.13	0.02	0	0	0.01
Planctomycetes	0.19	0.04	0.02	0.02	0.01	0.01	0.05	0
GN02	0	0	0.01	0.1	0	0	0	0.01
Tenericutes	0	0	0.06	0	0	0	0	0
Chloroflexi	0.05	0	0	0	0	0	0	0.01
WS3	0	0.04	0	0	0	0.02	0	0
Firmicutes	0.03	0	0	0	0.02	0.02	0	0.02
OD1	0	0	0	0.03	0	0.01	0	0.01
Fibrobacteres	0.02	0	0	0	0.01	0	0	0
Acidobacteria	0.02	0	0	0	0.02	0	0	0
Spirochaetes	0	0	0.02	0	0	0	0	0
Gemmatimonadetes	0	0	0.01	0	0	0	0	0
OP3	0.01	0	0	0	0.01	0.01	0	0
TM6	0	0	0.01	0	0	0	0	0
TM7	0	0	0	0	0.03	0.05	0	0.01
JS1	0	0	0	0	0.01	0	0	0
Zetaproteobacteria	0	0	0	0	0	0.01	0	0
Unclassified	0	0	0	0.02	0.03	0	0	0.01

Table 2-7. Sequencing statistics and diversity indices for bacterial communities

Tran-sect	Frontal zone	Station	Summary of SSU rRNA tags		Diversity indices			
			Total bacterial reads	No. of OTUs	Species richness		Species evenness	
					Chao1	ACE	Shannon	Simpson
1	SAZ	D-01	1191	82	105.88	116.16	3.48	0.05
	SAZ	D-25	1164	70	85.00	91.08	3.23	0.08
	SAZ	D-02	902	75	94.25	93.76	3.36	0.07
	SAZ	D-03	1113	76	103.00	106.28	3.24	0.10
	SAZ	D-04	1357	86	141.50	171.07	3.44	0.06
	SAZ	D-24	1007	100	141.63	141.63	3.78	0.04
	SAZ	D-05	1017	86	107.43	106.77	3.50	0.07
	SAZ	D-06	1246	83	106.21	107.26	3.38	0.09
	PFZ	D-07	1048	71	129.00	144.48	3.36	0.06
	PFZ	D-08	1215	65	80.83	87.76	3.28	0.07
	AZ	D-23	649	68	79.67	90.83	3.21	0.07
	AZ	D-09	1305	59	85.25	103.34	3.00	0.11
	AZ	D-10	1039	62	75.60	77.24	3.19	0.07
	AZ	D-21	1076	58	88.00	95.22	3.03	0.08
	AZ	D-22	1004	53	91.00	94.41	3.00	0.07
	AZ	D-11	994	46	63.14	82.26	2.85	0.08
	AZ	D-20	1025	55	68.13	68.97	2.96	0.09
	AZ	D-12	990	53	77.43	87.52	2.89	0.09
	AZ	D-19	1456	61	96.00	83.50	3.17	0.07
	AZ	D-17	1046	52	71.43	85.52	3.07	0.07
	AZ	D-13	970	54	78.43	99.99	3.12	0.06
	SZ	D-18	1267	61	91.00	83.66	3.16	0.07
	SZ	D-16	1288	52	90.00	92.24	2.95	0.08
	SZ	D-15	1165	53	80.14	103.42	3.10	0.06
	SZ	D-14	816	54	81.14	104.33	3.06	0.07

Table 2-7. To be continued

Tran-sect	Frontal zone	Station	Summary of SSU rRNA tags		Diversity indices			
			Total bacterial reads	No. of OTUs	Species richness		Species evenness	
					Chao1	ACE	Shannon	Simpson
	SAZ	T-31	1800	82	123.63	172.61	3.23	0.07
	SAZ	T-30	3486	90	142.11	246.66	3.22	0.08
	SAZ	T-29	1477	72	150.11	251.58	2.91	0.12
	SAZ	T-28	993	72	120.33	142.45	3.22	0.07
	SAZ	T-27	2493	74	115.33	119.95	3.22	0.07
	PFZ	T-26	1002	71	109.75	163.79	3.18	0.07
	PFZ	T-25	2240	74	116.27	146.14	3.29	0.06
	PFZ	T-24	3314	56	96.63	153.06	2.89	0.10
2	AZ	T-23	4068	62	134.50	206.69	3.11	0.07
	SPZ	T-22	4423	54	87.00	105.21	2.72	0.13
	SPZ	T-21	3886	49	91.00	96.36	2.40	0.17
	SPZ	T-20	3670	52	64.00	68.22	2.74	0.11
	SPZ	T-19	3879	71	107.91	142.81	3.09	0.09
	SPZ	T-18	3306	80	138.57	223.41	2.98	0.11
	SPZ	T-17	4164	62	94.50	132.00	2.89	0.10
	SPZ	T-16	3614	58	95.50	89.90	2.75	0.12
	SPZ	T-15	3575	54	81.60	116.04	2.54	0.15

In contrast, the abundance of *Cyanobacteria* and *Actinobacteria* diminished or disappeared in the south SAF (Table 2-6). *Cyanobacteria* composed 7.9% of the bacterial community at the northernmost site (D-01) of the SAZ, decreased to 0.6% at the southernmost site of the SAZ (D-06), and was not recovered at stations south of the SAF, suggesting SAF as a biogeographic boundary for this group (Fig. 2-7a and 2-8b). *Actinobacteria* accounted for less than 1.5% across all stations and showed a higher prevalence in the seawaters of the SAZ, with an abundance peak (1.4%) at D-03 and a relative abundance of 0.3% at stations located in the south SAF (Fig. 2-7a and Fig. 2-8b). The distribution of members of the *Verrucomicrobia* and SAR406 phyla was clearly associated with the PF. *Verrucomicrobia* showed a relative abundance of 6.8% at the lowest latitude (D-01), which decreased with increasing latitudes and was rarely detected at stations higher than 57°S near the PF ($r = –0.78$, $P < 0.001$; ($r = -0.78$, $P < 0.001$; Fig. 2-7a and Fig. 2-8b,c). Likewise, SAR406, which accounted for 0.3% to 2.5% north of the PF was almost not detected from the PFZ to the SZ. In transect 2, the SAR406 and *Verrucomicrobia* phyla showed significant negative correlations with latitude ($r = -0.79$ and -0.75, respectively, $P < 0.001$, Table 2-5). Both phyla had relative abundances less than 2% north of the SACCF and decreased as the latitude increased to become rarely detected or not detected south of SACCF (Fig. 2-7b and Fig. 2-8c). *Cyanobacteria* was not consistently recovered in the south of SAF (($r = -0.59$, $P < 0.01$; Fig. 2-7b and Fig. 2-8c).

Table 2-8. Pearson correlation coefficients between latitudinal change and diversity indices of bacterial communities

Diversity index	Latitude	
	Transect 1	**Transect 2**
No. of OTUs	-0.79***	-0.63**
Chao 1	-0.54**	-0.61**
ACE	-0.44*	-0.52**
Shannon	-0.70***	-0.65**
Simpson	0.13	0.56*

The statistical significance is represented as *, $p<0.05$, **, $p<0.01$, and ***, $p<0.001$

Relationship between bacterial community structure and environmental variables

Bacterial community structure was significantly distinguished by frontal zones (ANOSIM Global R=0.55, P<0.001) (Fig. 2-9). ANOSIM pairwise comparisons revealed that the Subpolar Zone (SPZ) was the most distinct among five frontal zones in terms of bacterial community compositions (0.92<R<1, P<0.01), and the SubAntarctic Zone (SAZ) were the next most distinct (0.80<R<0.91, P<0.01). No significant differences were observed between AZ and SZ in the bacterial community structures (all P > 0.05). When community compositions were compared between the same frontal zones in both transects, only SAZ significantly differed in both the bacterial and eukaryotic community structures (ANOSIM Global R=0.78, P<0.001).

Analysis of major OTUs

To investigate the presence of indicator species for each frontal zone among the dominant OTUs with relative abundances of >2%, the average relative abundance of each major OTU in each frontal zone was calculated and their shift patterns along the front was analyzed. Forty-three bacterial phylotypes belonged to *Cyanobacteria*, *Bacteroidetes*, SAR406, and *Proteobacteria* (Table 2-9). The relative abundances of some major OTUs clearly shifted according to the kind of front involved. Among the OTUs of *Bacteroidetes*, c2 with a 96.5% 16S rRNA sequence similarity with *Lacinutrix jangbogonensis* decreased, while c5 and c343 of the genus *Polaribacter* increased along the frontal zone related to latitudinal increase (Table 2-9). Some major OTUs such as c6, c9, and c35 of *Proteobacteria* showed a clear increase in abundance, while c116 and c14 decreased going from the SAZ to the SPZ (Table 2-9). Remarkably, the cyanobacterial OTU, c40 (96.1% 16S rRNA similarity with *Synechococcus rubescens*) and verrucomicrobial c79 (93.0% similarity with *Roseibacillus ponti*) showed dramatic decreases in abundance along the fronts. c40 showed an average abundance of 1.7% and 0.5%, respectively at the SAZ of

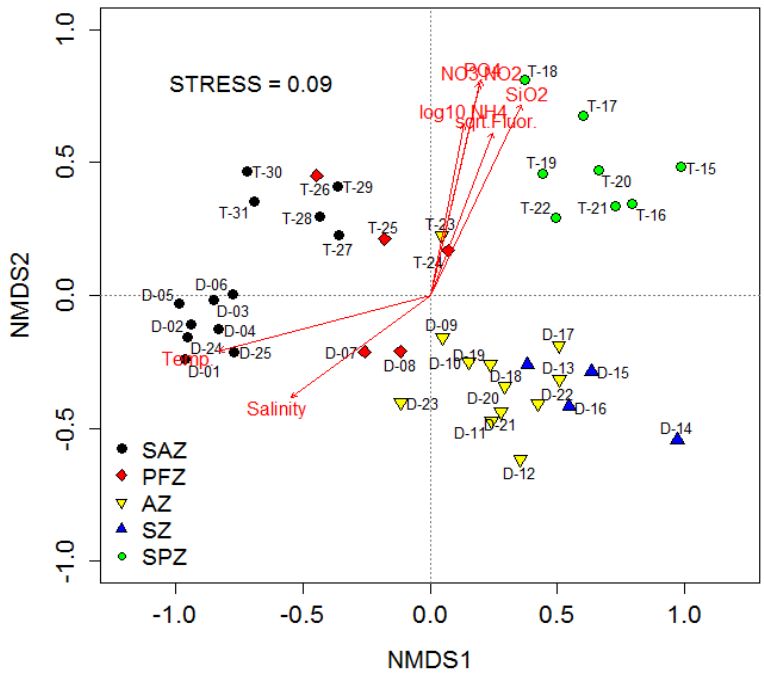


Figure 2-9. NMDS ordination plots of bacterial community structures based on Bray-Curtis dissimilarity matrix. Highly correlated and significant vectors of environmental variables (R^2>0.40, *P*<0.001) were overlaid onto the ordination space using the 'envfit' function in vegan R package. Abbreviations for each frontal zone are as follows: Subantarctic Zone (SAZ), Polar Frontal Zone (PFZ), the Antarctic Zone (AZ), Southern Zone (SZ), and Subpolar Zone (SZ) lies between the SACCF and SB, and Subpolar Zone (SPZ) is the south of the SB.

Table 2-9. Heat map and taxonomic affiliation of bacterial major OTUs according to the frontal zone

OTU_ID	Taxonomy						Relative abundance (%)							
							Transect 1				Transect 2			
	phylum	class	order	family	genus	species	SAZ	PFZ	AZ	SZ	SAZ	PFZ	AZ	SPZ
c146	Bacteroidetes	Flavobacteria	Flavobacteriales	Brumimicrobiaceae	EU800805_g	DQ009094_s	2.06	1.16	1.64	1.66	2.08	1.59	2.36	4.20
c43	Bacteroidetes	Flavobacteria	Flavobacteriales	Cryomorphaceae	DQ009083_g	FR683328_s	1.74	1.38	5.15	10.98	0.02	0.27	1.94	0.84
c66	Bacteroidetes	Flavobacteria	Flavobacteriales	EF572459_f	GQ259251_g	GQ259251_s	0.03	0.05	0.06	1.29	0.01	0.10	0.02	0.28
c67	Bacteroidetes	Flavobacteria	Flavobacteriales	EF572459_f	EF575089_g	EF575089_s	0.08	0.29	1.12	0.28	0.20	0.02	0.00	0.00
c20	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	ABCO_g	EU394565_s	1.30	2.83	9.18	9.65	5.72	10.52	8.60	1.48
c54	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	ABVV_g	HQ672042_s	12.58	1.72	0.15	0.00	2.11	1.15	0.25	0.02
c161	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	Chryseobacterium	Chryseobacterium aquaticum	1.04	0.00	1.06	0.00	0.00	0.00	0.00	0.00
c53	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	EF575148_g	DQ395894_s	1.87	3.42	2.56	2.75	4.62	4.28	4.25	6.30
c34	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	FR685471 g	FR685471 s	0.72	0.91	0.91	0.63	0.05	0.03	0.10	0.09
c8	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	FQ032819_g	AACY020239062_s	5.25	5.60	4.81	2.64	5.75	4.60	3.59	1.12
c2	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	Lacinutrix	AEYR_s	5.61	6.83	2.19	1.84	8.01	3.50	4.28	1.41
c5	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	Polaribacter	Polaribacter_irgensii	1.82	0.18	1.09	5.05	0.03	0.02	0.25	15.65
c11	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	Polaribacter	HQ730046_s	0.75	2.85	2.04	2.19	0.21	0.71	1.50	0.27
c343	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	Polaribacter	HQ730046_s	0.15	1.44	6.49	10.34	0.00	2.61	5.06	2.29
c22	Bacteroidetes	Flavobacteria	Flavobacteriales	GU061842_f	GU061842_g	HQ671900_s	4.26	3.00	1.52	0.91	2.32	3.38	1.97	0.15
c40	Cyanobacteria	Chroobacteria	Chroococcales	Prochlorococcaceae	Prochlorococcus	unclassified	1.74	0.00	0.00	0.00	0.45	0.00	0.00	0.00
c99	Proteobacteria	Betaproteobacteria	Burkholderiales	Derxia_f	DQ450176_g	AB240293_s	1.16	0.05	0.07	0.65	0.00	0.00	0.00	0.01
c24	Proteobacteria	Betaproteobacteria	Methylophilales	Methylophilaceae	OM43 g	HQ672174 s	0.21	0.69	0.82	0.69	0.47	0.60	0.98	0.43
c6	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	unclassified	unclassified	1.37	3.99	7.31	8.56	2.59	4.57	5.04	6.27
c7	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	Maritimibacter	Maritimibacter_alkaliphilus	0.90	3.90	3.94	2.58	1.30	1.42	1.06	0.77

c21	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	Loktanella	unclassified	0.28	0.00	0.47	1.33	0.03	0.01	0.02	1.45
c240	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	AVDB_g	EU799428_s	1.33	0.05	0.07	0.00	0.74	0.00	0.00	0.00
c27	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	HTCC2255_g	HTCC2255_s	2.96	0.27	0.87	1.54	0.71	0.32	0.20	1.90
c345	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	Octadecabacter	Octadecabacter_antarcticus	0.50	0.00	0.22	1.50	0.00	0.00	0.00	0.85
c29	Proteobacteria	Alphaproteobacteria	Rhizobiales	EF574506_f	EU805215_g	EU805215_s	0.63	0.36	0.08	0.00	0.77	0.59	0.07	0.06
c1	Proteobacteria	Alphaproteobacteria	SAR11	SAR11-1_f	Pelagibacter	unclassified	1.09	0.84	0.63	0.22	1.32	1.24	0.76	0.78
c13	Proteobacteria	Alphaproteobacteria	SAR11	SAR11-1_f	Pelagibacter	Pelagibacter_ubique	10.58	17.34	18.38	10.36	20.23	19.07	18.09	21.60
c116	Proteobacteria	Alphaproteobacteria	SAR11	SAR11-2_f	EU801223_g	AF353223_s	2.12	2.49	1.97	0.41	3.59	3.77	3.42	1.13
c16	Proteobacteria	Alphaproteobacteria	SAR11	HQ671838_f	HQ671838_g	HQ671838_s	2.51	0.18	0.22	0.08	0.81	0.27	0.32	0.17
c26	Proteobacteria	Alphaproteobacteria	SAR116	OM38 f	OM38 g	EF471703 s	0.47	0.86	1.18	0.71	1.06	0.50	0.25	0.41
c59	Proteobacteria	Gammaproteobacteria	Alteromonadales	Alteromonadaceae	Alteromonas	Alteromonas_marina	2.04	1.62	0.01	0.04	0.02	0.00	0.00	0.02
c23	Proteobacteria	Gammaproteobacteria	Alteromonadales	Colwelliaceae	Colwellia	unclassified	0.59	0.00	0.16	0.08	0.05	0.12	0.10	0.64
c108	Proteobacteria	Gammaproteobacteria	Alteromonadales	Marinobacter_f	Marinobacter	Marinobacter_algicola	0.09	4.58	0.00	0.00	0.00	0.00	0.00	0.00
c279	Proteobacteria	Gammaproteobacteria	Alteromonadales	Pseudoalteromonadaceae	Pseudoalteromonas	unclassified	1.02	1.28	0.01	0.00	0.00	0.00	0.00	0.01
c71	Proteobacteria	Gammaproteobacteria	Alteromonadales	Pseudoalteromonadaceae	Pseudoalteromonas	Pseudoalteromonas_distincta	0.50	3.76	0.19	0.02	0.04	0.00	0.02	0.03
c9	Proteobacteria	Gammaproteobacteria	Alteromonadales	Porticoccus_f	SAR92_g	unclassified	0.55	1.85	2.99	4.37	1.39	2.81	3.66	2.80
c35	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Oceanospirillaceae	AM402959_g	AM402959_s	0.19	0.35	1.23	2.16	0.21	3.04	6.83	10.68
c344	Proteobacteria	Gammaproteobacteria	Ruthia_o	Ruthia_f	Ruthia	GU235569_s	0.50	1.34	1.06	1.04	1.21	3.08	3.86	3.10
c4	Proteobacteria	Gammaproteobacteria	SAR86	SAR86 f	SAR86 g	unclassified	3.02	5.82	8.21	5.42	11.49	8.73	7.99	3.42
c14	Proteobacteria	Gammaproteobacteria	SAR86	SAR86_f	SAR86_g	OM23_s	2.13	3.37	1.11	0.16	4.36	4.06	2.85	1.06
c154	SAR406	SAR406_c	SAR406_o	SAR406_f	AQSA_g	AQSA_s	2.23	0.19	0.00	0.00	0.27	0.00	0.00	0.00

Major OTUs with 2% or higher compositions at least in one sample were selected. The color gradient from white to black indicates the lowest to highest relative abundance values [§] denotes taxonomic information retrieved by EzTaxon-e database search (Kim et al.., 2012). [†] represents the average value of relative abundance at each frontal zone.

both transects was not recovered at the south SAF. The gradual decrease of c40 along the latitude of transect 1 is remarkable. c40, which showed a relative abundance of 7.9% at station D-01 in the SAZ gradually decreased up to 0.6% at D-06 (55°S), the approximately southern boundary of the SAF and then it was not recovered in the surface water of stations higher than SAF indicating that it is an indicator species for the SAZ (Table 2-10). c79 of *Verrucomicrobia* approximately a 1% abundance in the SAZ was rarely detected or not detected in the southern zones of the PF (Table 2-9). In contrast, c13 with 99.3% 16S rRNA sequence similarity with *Pelagibacter ubique* accounted for 7.0% to 29.0% of the bacterial community (more than 10% at all stations except for 2 stations, D-13 and D-14) throughout the all stations in transect 1 and 2.

Table 2-10. Heat map and taxonomic affiliation of bacterial major OTUs. Major OTUs with 2% or higher compositions at least in one sample were selected. The color gradient from white to black indicates the lowest to highest relative abundance values

OTU_ID	Taxonomy: phylum	class	order	family	genus	species
c2	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	Lacinutrix	AEYR_s
c8	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	FQ032819_g	AACY020239062_s
c22	Bacteroidetes	Flavobacteria	Flavobacteriales	GU061842_f	GU061842_g	HQ671900_s
c5	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	Polaribacter	Polaribacter_irgensii
c53	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	EF575148_g	DQ395894_s
c146	Bacteroidetes	Flavobacteria	Flavobacteriales	Brumimicrobiaceae	EU800805_g	DQ009094_s
c1	Proteobacteria	Alphaproteobacteria	SAR11	SAR11-1_f	Pelagibacter	unclassified
c6	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	unclassified	unclassified
c11	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	Polaribacter	HQ730046_s
c7	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	Maritimibacter	Maritimibacter_alkaliphilus
c14	Proteobacteria	Gammaproteobacteria	SAR86	SAR86_f	SAR86_g	OM23_s
c16	Proteobacteria	Alphaproteobacteria	SAR11	HQ671838_f	HQ671838_g	HQ671838_s
c20	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	ABCO_g	EU394565_s
c21	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	Loktanella	unclassified
c9	Proteobacteria	Gammaproteobacteria	Alteromonadales	Porticoccus_f	SAR92_g	unclassified
c23	Proteobacteria	Gammaproteobacteria	Alteromonadales	Colwelliaceae	Colwellia	unclassified
c24	Proteobacteria	Betaproteobacteria	Methylophilales	Methylophilaceae	OM43_g	HQ672174_s
c4	Proteobacteria	Gammaproteobacteria	SAR86	SAR86_f	SAR86_g	unclassified
c27	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	HTCC2255_g	HTCC2255_s
c29	Proteobacteria	Alphaproteobacteria	Rhizobiales	EF574506_f	EU805215_g	EU805215_s
c34	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	FR685471_g	FR685471_s
c35	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Oceanospirillaceae	AM402959_g	AM402959_s
c40	Cyanobacteria	Chroobacteria	Chroococcales	Prochlorococcaceae	Prochlorococcus	unclassified
c43	Bacteroidetes	Flavobacteria	Flavobacteriales	Cryomorphaceae	DQ009083_g	FR683328_s
c13	Proteobacteria	Alphaproteobacteria	SAR11	SAR11-1_f	Pelagibacter	Pelagibacter_ubique
c54	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	ABVV_g	HQ672042_s
c59	Proteobacteria	Gammaproteobacteria	Alteromonadales	Alteromonadaceae	Alteromonas	Alteromonas_marina
c66	Bacteroidetes	Flavobacteria	Flavobacteriales	EF572459_f	GQ259251_g	GQ259251_s
c67	Bacteroidetes	Flavobacteria	Flavobacteriales	EF572459_f	EF575089_g	EF575089_s
c71	Proteobacteria	Gammaproteobacteria	Alteromonadales	Pseudoalteromonadaceae	Pseudoalteromonas	Pseudoalteromonas_distincta
c79	Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	Roseibacillus_f	CAGM_g	DQ395400_s
c99	Proteobacteria	Betaproteobacteria	Burkholderiales	Derxia_f	DQ450176_g	AB240293_s
c108	Proteobacteria	Gammaproteobacteria	Alteromonadales	Marinobacter_f	Marinobacter	Marinobacter_algicola
c26	Proteobacteria	Alphaproteobacteria	SAR116	OM38_f	OM38_g	EF471703_s
c116	Proteobacteria	Alphaproteobacteria	SAR11	SAR11-2_f	EU801223_g	AF353223_s
c154	SAR406	SAR406_c	SAR406_o	SAR406_f	AQSA_g	AQSA_s
c161	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	Chryseobacterium	Chryseobacterium_aquaticum
c240	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	AVDB_g	EU799428_s
c279	Proteobacteria	Gammaproteobacteria	Alteromonadales	Pseudoalteromonadaceae	Pseudoalteromonas	unclassified
c343	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	Polaribacter	HQ730046_s
c344	Proteobacteria	Gammaproteobacteria	Ruthia_o	Ruthia_f	Ruthia	GU235569_s
c345	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	Octadecabacter	Octadecabacter_antarcticus

OTU_ID	Transect 1: SAZ								PFZ		AZ											SZ				Average	r
	D-01	D-25	D-02	D-03	D-04	D-24	D-05	D-06	D-07	D-08	D-23	D-09	D-10	D-21	D-22	D-11	D-20	D-12	D-19	D-17	D-13	D-18	D-16	D-15	D-14		
c2	10.24	5.76	9.20	4.49	4.13	3.08	1.77	6.18	5.44	8.23	8.47	1.76	2.12	1.30	1.99	1.91	1.17	1.01	1.92	1.34	1.13	1.89	1.48	2.15	1.84	3.60	-0.73
c8	3.86	2.66	3.44	5.66	6.56	4.77	6.88	5.30	4.87	6.34	6.32	4.60	7.60	5.95	5.98	6.24	3.61	1.52	3.91	3.54	3.61	3.00	4.04	1.80	1.72	4.55	-0.21
c22	2.18	2.66	2.66	3.50	3.68	3.67	5.01	5.62	2.96	3.05	2.47	1.30	2.41	2.04	1.69	0.60	0.88	0.81	2.20	1.24	1.03	1.50	1.48	0.43	0.25	2.21	-0.64
c5	0.00	0.00	0.00	0.00	0.15	0.00	0.00	0.00	0.29	0.08	0.00	0.00	0.58	0.84	1.20	0.30	0.49	0.00	0.48	3.06	5.05	1.58	1.86	6.09	10.66	1.31	0.53
c53	3.95	5.76	3.22	1.35	5.45	3.48	2.85	3.77	3.63	3.21	2.77	2.22	2.21	1.67	3.09	3.12	2.83	2.53	2.40	3.44	1.86	2.53	3.88	2.75	1.84	3.03	-0.50
c146	1.01	0.09	0.33	0.36	0.29	0.10	0.79	0.24	0.67	1.65	0.77	2.30	2.02	1.39	1.89	1.51	2.73	0.61	0.76	1.53	2.58	2.37	0.70	1.63	1.96	1.21	0.65
c1	1.85	1.89	1.66	2.52	1.25	1.49	1.97	1.69	0.86	0.82	0.92	1.00	0.38	0.74	0.10	0.50	0.68	0.71	0.76	0.67	0.52	0.32	0.31	0.26	0.00	0.95	-0.89
c6	0.25	0.86	1.22	0.72	0.81	0.70	1.18	1.93	2.96	5.02	5.70	4.37	8.66	8.09	9.06	4.53	8.10	7.37	5.56	8.70	10.31	8.13	6.68	9.01	10.42	5.21	0.91
c11	0.59	1.72	0.55	0.99	0.22	0.10	0.10	0.00	1.91	3.79	1.54	2.84	1.25	0.84	0.60	1.81	1.76	1.72	6.25	1.72	2.16	2.92	0.47	2.58	2.82	1.65	0.45
c7	0.25	1.29	0.89	1.08	0.96	2.58	1.38	1.77	4.10	3.70	5.08	2.68	5.20	3.35	1.99	1.11	6.24	3.54	7.21	3.82	3.09	3.55	2.48	3.18	1.10	2.87	0.54
c14	2.77	4.98	2.66	3.23	3.10	2.78	2.16	2.17	2.39	4.36	2.31	2.76	1.06	0.19	0.60	0.40	1.07	0.71	1.65	0.38	1.03	0.47	0.08	0.09	0.00	1.74	-0.81
c16	3.02	2.84	2.44	0.99	1.03	1.49	2.16	2.41	0.19	0.16	0.31	0.61	0.38	0.37	0.00	0.10	0.20	0.10	0.00	0.10	0.21	0.32	0.00	0.00	0.00	0.78	-0.87
c20	1.43	0.00	0.55	0.81	0.15	0.10	0.20	1.20	2.19	3.46	2.62	5.52	7.70	9.67	14.94	10.76	9.66	11.82	8.59	10.99	8.76	15.31	10.25	5.92	7.11	5.99	0.82
c21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.09	0.30	0.00	0.29	0.00	0.27	2.20	1.96	1.34	0.70	1.20	2.08	0.42	0.58
c9	0.84	1.46	1.33	1.08	1.47	1.19	0.59	0.80	1.15	2.55	1.54	2.76	1.83	4.09	2.39	4.63	1.85	4.55	1.65	2.77	4.85	2.29	2.17	6.52	6.50	2.51	0.67
c23	0.00	0.09	0.00	0.00	0.22	0.30	0.00	0.00	0.00	0.00	0.00	0.00	0.29	0.28	0.30	0.00	0.29	0.00	0.14	0.38	0.10	0.08	0.23	0.00	0.00	0.11	0.25
c24	0.00	0.09	0.33	0.81	0.29	0.10	0.59	0.48	1.05	0.33	0.31	1.53	2.31	0.56	0.30	0.30	1.66	0.20	0.76	0.38	0.72	0.95	1.24	0.43	0.12	0.63	0.31
c4	5.71	7.39	4.88	5.39	5.23	4.77	7.08	5.22	5.63	6.01	7.24	9.58	7.60	6.23	8.76	9.36	12.68	8.99	7.01	5.54	7.32	4.89	9.47	5.24	2.08	6.77	0.24
c27	0.25	0.34	0.22	0.18	0.07	0.00	0.20	0.40	0.38	0.16	0.00	0.92	1.06	0.46	0.90	0.70	0.88	1.52	0.48	1.05	1.55	1.03	1.16	2.75	1.23	0.72	0.69
c29	0.25	1.12	1.88	0.18	0.74	2.18	0.98	1.04	0.48	0.25	0.15	0.46	0.10	0.00	0.00	0.00	0.00	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.40	-0.68
c34	2.18	0.09	0.33	0.00	0.22	0.20	0.00	0.08	0.67	1.15	0.77	0.92	1.35	0.37	0.20	1.71	2.24	0.10	1.58	0.19	0.62	1.34	0.85	0.34	0.00	0.70	0.12
c35	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.29	0.41	0.15	1.23	1.83	1.39	0.90	0.60	0.88	1.31	1.30	1.91	2.06	1.74	1.79	2.92	2.21	0.92	0.83
c40	7.89	1.80	3.22	3.77	3.98	4.47	2.06	0.64	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.11	-0.82
c43	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.95	1.81	1.39	2.45	1.44	5.48	10.56	10.46	2.24	4.95	2.82	7.27	7.63	6.00	13.74	12.19	12.01	4.14	0.76
c13	15.20	23.28	23.28	26.42	21.22	15.59	20.94	23.27	15.17	19.51	18.64	29.04	19.25	19.52	9.96	17.00	21.76	18.99	17.24	16.54	14.23	14.05	11.72	8.67	6.99	17.90	-0.52
c54	3.95	2.41	6.10	3.41	4.72	4.67	3.24	3.61	1.62	1.81	0.92	0.46	0.10	0.00	0.00	0.10	0.00	0.00	0.07	0.00	0.00	0.00	0.00	0.00	0.00	1.49	-0.89
c59	0.00	0.09	0.00	0.00	0.00	0.00	0.39	0.00	3.24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.00	0.00	0.00	0.00	0.16	0.00	0.00	0.16	-0.08
c66	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.62	0.32	0.54	1.97	2.33	0.24	0.44
c67	0.00	0.17	0.00	0.00	0.07	0.30	0.79	0.00	0.10	0.49	0.00	0.00	0.48	1.58	1.59	3.12	1.37	1.62	1.44	0.29	0.82	0.63	0.39	0.09	0.00	0.61	0.43
c71	0.17	0.26	0.55	0.36	5.16	0.10	0.00	0.00	7.44	0.08	0.46	0.00	0.29	0.19	0.00	0.00	0.10	0.91	0.14	0.00	0.00	0.00	0.00	0.09	0.00	0.65	-0.21
c79	4.45	0.69	1.00	2.88	1.47	2.58	1.87	2.33	0.86	0.41	0.00	0.08	0.00	0.00	0.00	0.00	0.10	0.00	0.00	0.00	0.00	0.16	0.00	0.00	0.00	0.75	-0.79
c99	0.00	0.00	0.00	0.00	0.07	1.29	0.00	0.00	0.10	0.00	0.00	0.00	0.00	0.09	0.00	0.00	0.00	0.00	0.62	0.00	0.10	2.05	0.00	0.43	0.12	0.20	0.17
c108	0.00	0.00	0.00	0.00	0.00	0.10	0.00	0.00	9.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.37	-0.07
c26	0.76	0.34	0.67	1.53	0.96	0.60	1.08	1.52	0.57	1.15	3.08	0.84	2.50	1.12	1.49	0.80	1.07	0.30	0.55	0.67	0.52	0.79	0.93	0.52	0.61	1.00	-0.01
c116	1.60	4.90	2.88	3.86	4.27	3.08	3.05	2.81	2.67	2.30	4.16	3.45	1.54	2.23	1.20	1.51	1.56	1.31	1.85	1.05	1.86	0.71	0.47	0.34	0.12	2.19	-0.71
c154	0.34	0.69	0.44	2.43	0.96	0.99	2.36	1.04	0.38	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.39	-0.61
c161	0.00	3.26	0.00	0.00	0.00	4.17	0.00	0.00	0.00	0.00	8.01	0.00	0.00	1.30	1.69	0.00	0.49	0.00	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.76	-0.15
c240	2.02	1.29	2.99	1.80	0.88	2.68	1.38	0.80	0.10	0.00	0.77	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.59	-0.85
c279	0.08	0.00	0.00	0.00	2.28	0.00	0.00	0.08	2.39	0.16	0.00	0.00	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20	-0.21
c343	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.24	1.65	0.31	2.22	3.37	7.90	8.07	10.46	2.34	13.94	7.14	8.41	7.22	8.68	10.48	11.16	11.03	4.62	0.81
c344	0.08	2.06	0.89	0.54	1.11	1.19	1.28	0.80	2.10	0.58	0.15	2.38	0.96	0.84	1.10	0.80	1.07	0.91	1.30	1.72	0.41	0.55	2.25	1.37	0.00	1.06	0.03
c345	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.96	1.44	0.32	0.78	2.23	2.70	0.34	0.50

OTU_ID	Transect 2: SAZ					PFZ			AZ	SPZ								Average	r
	T-31	T-30	T-29	T-28	T-27	T-26	T-25	T-24	T-23	T-22	T-21	T-20	T-19	T-18	T-17	T-16	T-15		
c2	9.44	6.86	9.41	11.18	3.17	4.49	3.39	2.63	4.28	1.67	0.82	2.26	1.42	1.00	2.02	1.49	0.56	3.89	-0.85
c8	5.67	4.45	5.75	6.34	6.54	5.59	5.98	2.23	3.59	1.13	0.33	0.57	2.71	1.36	2.09	0.36	0.42	3.24	-0.84
c22	2.33	2.15	2.17	1.81	3.13	5.39	2.81	1.93	1.97	0.47	0.08	0.03	0.13	0.42	0.05	0.00	0.00	1.46	-0.73
c5	0.00	0.00	0.07	0.10	0.00	0.00	0.00	0.06	0.25	2.94	27.23	14.14	4.20	3.12	23.13	21.64	28.84	7.39	0.72
c53	8.33	5.88	2.98	2.62	3.29	4.79	4.29	3.77	4.25	2.35	1.52	1.69	21.22	12.07	10.81	0.47	0.28	5.33	0.13
c146	1.78	1.06	1.90	2.42	3.25	1.40	1.34	2.02	2.36	3.62	2.75	9.13	4.87	4.14	4.49	2.91	1.71	3.01	0.48
c1	0.89	1.89	1.62	0.40	1.81	1.60	0.94	1.18	0.76	0.63	0.93	0.95	0.52	1.39	0.43	0.42	0.98	1.02	-0.50
c6	1.56	2.55	2.10	1.91	4.81	3.59	5.36	4.77	5.04	8.50	3.29	5.42	8.17	4.90	4.18	8.83	6.88	4.82	0.74
c11	0.56	0.23	0.20	0.00	0.08	0.30	1.65	0.18	1.50	0.05	0.18	0.33	0.52	0.24	0.48	0.14	0.20	0.40	-0.04
c7	0.56	0.55	2.30	2.42	0.68	2.30	1.03	0.94	1.06	0.38	0.62	1.12	0.31	0.51	0.26	1.72	1.26	1.06	-0.27
c14	3.72	4.76	2.10	5.84	5.38	4.69	4.15	3.35	2.85	0.88	0.80	0.38	1.06	3.06	1.42	0.50	0.39	2.67	-0.75
c16	0.44	1.15	0.20	1.91	0.36	0.50	0.18	0.12	0.32	0.11	0.23	0.41	0.21	0.12	0.12	0.08	0.06	0.38	-0.55
c20	8.89	7.37	5.28	1.11	5.94	8.78	11.12	11.65	8.60	4.88	0.85	0.95	1.75	0.18	1.25	0.77	1.20	4.74	-0.62
c21	0.00	0.03	0.07	0.00	0.04	0.00	0.00	0.03	0.02	0.18	1.21	0.30	0.26	0.30	0.86	3.82	4.67	0.69	0.62
c9	0.94	0.63	1.56	2.11	1.72	1.00	2.77	4.68	3.66	2.74	2.62	1.93	2.73	1.91	4.35	4.04	2.10	2.44	0.57
c23	0.00	0.11	0.00	0.00	0.12	0.10	0.22	0.03	0.10	0.09	0.28	0.76	2.19	1.24	0.19	0.14	0.22	0.34	0.46
c24	0.00	0.11	0.68	0.81	0.76	0.40	0.67	0.72	0.98	0.70	0.33	0.27	0.41	0.64	0.43	0.39	0.28	0.51	-0.01
c4	10.89	10.33	12.19	13.70	10.35	9.28	8.44	8.48	7.99	3.57	2.88	2.75	4.12	2.81	4.06	3.02	4.17	7.00	-0.90
c27	0.22	0.29	0.81	1.01	1.24	0.60	0.00	0.36	0.20	0.90	0.93	2.72	0.98	0.33	0.53	4.10	4.73	1.17	0.56
c29	0.22	0.92	0.14	1.21	1.36	1.30	0.22	0.24	0.07	0.00	0.00	0.00	0.13	0.12	0.02	0.06	0.14	0.36	-0.57
c34	0.11	0.00	0.14	0.00	0.00	0.00	0.00	0.09	0.10	0.02	0.10	0.00	0.08	0.00	0.17	0.36	0.03	0.07	0.31
c35	0.00	0.11	0.07	0.10	0.76	0.80	3.97	4.35	6.83	28.83	12.07	15.97	5.54	6.35	4.80	7.75	4.11	6.02	0.46
c40	0.72	1.26	0.07	0.10	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	-0.60
c43	0.00	0.00	0.00	0.00	0.12	0.00	0.31	0.48	1.94	1.31	0.80	0.57	1.26	0.21	0.48	1.11	0.98	0.56	0.60
c13	14.17	21.20	27.83	16.72	21.26	19.46	15.18	22.57	18.09	19.01	27.05	20.68	16.76	25.32	15.13	24.05	24.84	20.55	0.21
c54	3.78	2.90	0.41	1.81	1.64	2.30	1.12	0.03	0.25	0.00	0.00	0.00	0.00	0.09	0.02	0.03	0.00	0.85	-0.81
c59	0.11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.03	0.05	0.00	0.00	0.00	0.00	0.03	0.01	-0.18
c66	0.06	0.00	0.00	0.00	0.00	0.30	0.00	0.00	0.02	0.20	0.90	0.93	0.05	0.06	0.00	0.03	0.08	0.15	0.23
c67	0.33	0.14	0.20	0.30	0.04	0.00	0.04	0.03	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.07	-0.76
c71	0.00	0.00	0.07	0.00	0.12	0.00	0.00	0.00	0.02	0.00	0.00	0.03	0.00	0.18	0.02	0.00	0.00	0.03	0.05
c79	0.72	1.06	0.20	1.21	1.89	2.00	0.71	0.24	0.39	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50	-0.66
c99	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.00	0.00	0.00	0.29
c108	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-
c26	1.06	1.20	1.08	1.31	0.64	0.80	0.40	0.30	0.25	0.18	0.10	0.35	0.93	0.76	0.29	0.39	0.25	0.61	-0.66
c116	3.11	3.16	4.94	3.32	3.41	2.79	5.04	3.47	3.42	1.47	0.87	1.63	1.11	1.66	0.82	0.94	0.53	2.45	-0.79
c154	0.56	0.43	0.00	0.30	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.08	-0.66
c161	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-
c240	1.61	0.37	1.02	0.40	0.28	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.22	-
c279	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.00	0.00	0.00	0.01	0.26
c343	0.00	0.00	0.00	0.00	0.00	0.00	0.54	7.30	5.06	4.02	1.98	2.21	3.58	0.27	3.99	0.86	1.40	1.84	0.38
c344	0.67	1.38	0.27	2.01	1.72	1.90	3.57	3.77	3.86	1.61	1.52	0.84	3.35	11.92	4.08	0.72	0.73	2.58	0.31
c345	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.14	2.08	1.53	0.26	0.39	0.14	1.16	1.06	0.40	0.57

2.3.2 Vertical and horizontal distribution of bacterial communities in seawater of the Ross Sea

Change of environmental factors

There were obvious changes in environmental parameters measured in the upper layers of the water column, while little variation was observed below 150 m at all stations, indicating a strong stratification within a water column (Fig. 2-4 and Table 2-1). Temperature, chlorophyll (Chl, measured by evaluating the fluorescence intensity), and oxygen concentration decreased, whereas salinity, concentration of inorganic nutrients (phosphate, nitrate, nitrite, and silicate), and pressure increased as depth increased (Fig. 2-4a–d and Table 2-1). The average salinities were 33.6, 34.1, and 34.6 psi in the mixed, thermocline, and deep layers, respectively. The seawater temperature showed a trend opposite that of salinity, dropping from -0.54°C in the surface water to -1.8°C in deeper waters.

Cluster analysis based on the inorganic nutrient concentration, oxygen concentration, temperature, salinity, fluorescence, and pressure by Euclidean distance showed that samples deeper than 400 m were clearly distinct from other samples. Moreover, the remaining samples were largely clustered according to the water layer, regardless of the sampling location (Fig. 2-10). On the basis of vertical variations in environmental parameters, the water layers were divided into the mixed layer, thermocline layer, and deep waters (Table 2-1).

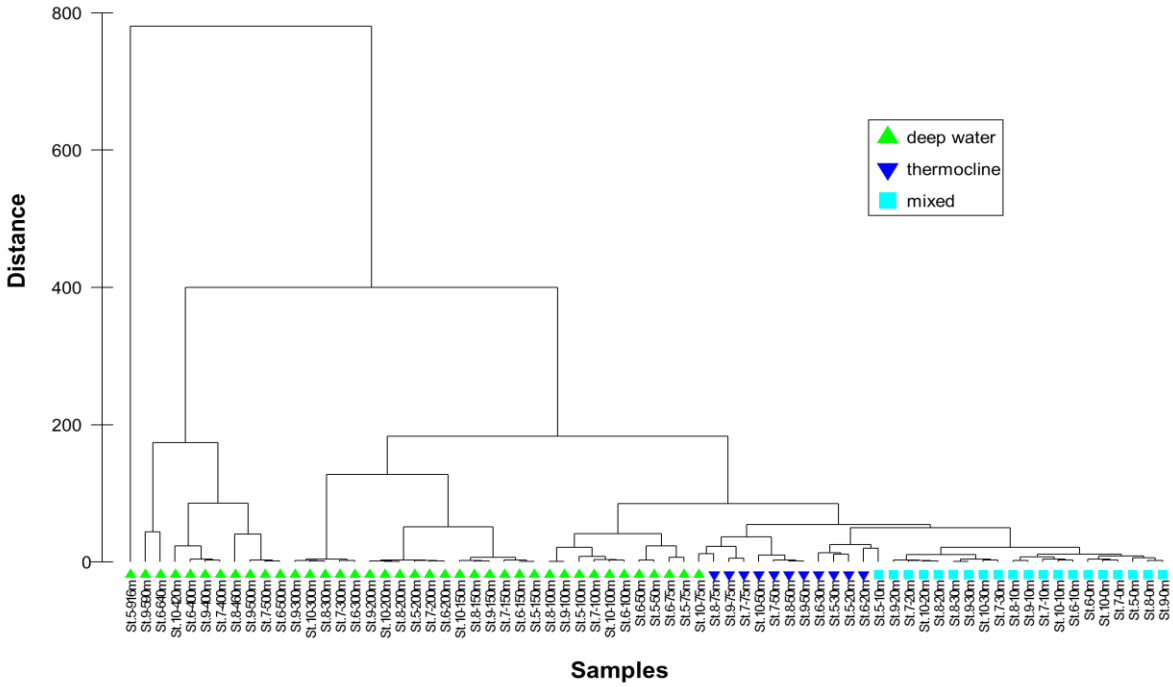


Figure 2-10. Similarities among samples according to environmental variables. Environmental variables included inorganic compound concentrations (phosphate, nitrite, nitrate, silicate, and ammonium), temperature, salinity, fluorescence, oxygen concentration, and pressure.

Vertical bacterial community structure

Diversity indices, such as number of OTUs and ACE, Shannon, and Simpson indices, increased with depth (Tables 2-11 and 2-12). Out of the 27 bacterial phyla or classes of *Proteobacteria* recovered, *Bacteroidetes*, *Alphaproteobacteria*, and *Gammaproteobacteria* dominated, making up 57.3–99.8% of the bacterial community, throughout the whole water column at all stations (Figs. 2-11 and 2-12). Analysis of the surface samples (i.e., those with a depth of 30 m or less) showed that over 98.1% of the abundance at all stations was dominated by three groups: *Bacteroidetes* (52.0% ± 6.6%), *Alphaproteobacteria* (30.1% ± 2.8%), and *Gammaproteobacteria* (17.0% ± 5.3%; Table S4). In the thermocline zone, the total abundance of the three dominant groups was 96.6–99.8% (*Bacteroidetes* [48.0% ± 7.1%], *Alphaproteobacteria* [21.3% ± 5.2%], and *Gammaproteobacteria* [29.1% ± 2.9%]), ranging from 57.3% to 98.3% in the deep-zone communities (24.3% of *Bacteroidetes* [SD ± 4.3%], 33.8% of *Alphaproteobacteria* [SD ± 6.1%], and 30.4% of *Gammaproteobacteria* [SD ± 5.2%]; Fig. 2-12A and Table 2-13). When the relative abundance was calculated according to the water layer at each station, the minor phyla *Chloroflexi*, *Gemmatimonadetes*, *Planctomycetes*, and *Firmicutes* accounted for less than 2.5% of the total community and rarely recovered or did not recover in the mixed layer and thermocline zone, with greater abundance observed in deeper layers (Fig. 2-11 and Table 2-12). As the depth increased, the relative abundance of *Bacteroidetes* decreased, whereas those of Gammaproteobacteria, Deltaproteobacteria, SAR406, Verrucomicrobia, Actinobacteria, *Chloroflexi*, and *Gemmatimonadetes* increased in at least six stations and those of other phyla, such as *Planctomycetes* and *Lentisphaerae*, also increased in at least four stations (Table 2-14).

Table 2-11. Diversity indices

Sample ID	Total reads	No. of OTUs	ACE	Shannon	Simpson
Z1	1375	31	42.19	2.81	11.78
Z2	1116	33.6	49.31	2.88	12.37
Z4	835	31	41.14	2.79	10.88
Z3	1033	32.5	46.04	2.84	11.96
Z5	1057	36.3	51.97	2.93	12.35
Z6	850	32.3	45.56	2.84	11.82
Z8	1079	31.9	49.02	2.84	12.55
Z9	1384	32.9	54.59	2.79	11.19
Z10	1176	34.5	46.23	2.95	13.34
Z7	980	38	59.26	2.93	12.03
St.1-0m	988	36.5	53.86	2.92	12.26
St.1-10m	976	38.3	62.96	2.97	13.03
St.1-20m	1543	41.1	60.90	3.06	13.63
St.1-30m	1319	44.1	62.34	3.15	14.22
St.1-50m	876	42.1	57.01	3.13	14.54
St.1-75m	1256	49.3	83.20	3.20	14.60
St.1-100m	638	39.1	53.21	3.11	15.17
St.1-150m	1256	41.7	69.66	3.10	15.23
St.2-0m	789	34.4	43.79	2.94	12.59
St.2-10m	910	38.5	55.83	2.97	12.13
St.2-20m	652	37.7	55.93	2.92	11.04
St.2-30m	1052	45.5	63.76	3.10	11.56
St.2-50m	998	44	58.92	3.04	10.56
St.2-75m	1043	44.2	74.26	3.01	11.36
St.2-100m	1125	45.7	76.71	3.18	15.58
St.2-150m	832	39.7	61.81	3.07	14.68
St.2-200m	1276	49.4	93.43	3.21	14.93
St.3-0m	957	30.1	46.24	2.73	10.75
St.3-10m	1670	31.9	44.60	2.81	11.05
St.3-20m	1282	30.5	38.39	2.76	10.74
St.3-30m	1087	38.1	52.37	3.10	15.94
St.3-50m	983	40.7	52.67	3.16	15.87
St.3-75m	1073	44.2	68.60	3.19	15.70
St.3-100m	2140	39.3	60.70	3.06	14.11
St.3-150m	1525	49	99.99	3.18	14.59
St.3-200m	596	47.3	80.89	3.27	17.88
St.3-300m	189	nd	nd	nd	nd
St.3-445m	41	nd	nd	nd	nd

Table 2-11. To be continued

Sample ID	Total reads	No. of OTUs	ACE	Shannon	Simpson
St.4-0m	1045	35	49.86	2.87	11.00
St.4-10m	1544	39	58.96	3.00	12.74
St.4-20m	1014	29.9	38.41	2.76	10.73
St.4-30m	955	33.2	48.30	2.79	10.59
St.4-50m	1256	43.3	58.66	3.13	12.98
St.4-75m	395	40.8	58.16	3.06	13.46
St.4-100m	1102	39.7	64.17	3.02	13.32
St.4-150m	835	36	65.41	2.90	12.53
St.4-200m	1201	44.1	75.15	3.12	14.18
St.4-300m	825	41.6	65.72	3.10	14.94
St.4-335m	1094	51.5	92.34	3.30	16.44
St.10-0m	1248	36.7	55.63	2.81	9.95
St.10-10m	1887	36.3	51.80	2.87	10.69
St.10-20m	1391	36.9	47.47	2.93	11.40
St.10-30m	1000	38.1	50.70	2.97	11.28
St.10-50m	1420	38.1	52.77	3.06	14.18
St.10-75m	537	43.8	59.51	3.21	16.29
St.10-100m	345	43	57.28	3.26	18.38
St.10-150m	358	38.2	61.40	3.07	15.44
St.10-200m	452	44.3	66.65	3.18	14.96
St.10-300m	122	nd	nd	nd	nd
St.10-420m	201	nd	nd	nd	nd
St.8-0m	2338	41.3	63.79	2.88	9.88
St.8-10m	2420	39.9	53.31	2.86	9.47
St.8-20m	1934	39.4	50.69	2.95	10.69
St.8-30m	980	31	40.59	2.58	7.57
St.8-50m	1061	32	46.96	2.74	9.55
St.8-75m	1105	40.9	53.06	3.05	12.27
St.8-100m	1100	44.7	78.82	3.00	10.61
St.8-150m	1092	40.4	57.79	3.14	16.28
St.8-200m	1136	40.1	73.40	3.03	14.57
St.8-300m	1148	47.1	91.16	2.78	6.43
St.8-460m	772	53.3	90.92	3.35	17.15
St.9-0m	891	34.9	45.51	2.93	11.92
St.9-10m	1097	37.5	49.23	2.87	9.82
St.9-20m	992	35.7	46.33	2.87	10.12
St.9-30m	1137	37.5	47.65	3.00	12.70
St.9-50m	680	37.1	56.20	2.85	10.09
St.9-75m	878	42	56.16	3.19	16.56
St.9-100m	1039	43	61.82	3.22	17.17
St.9-150m	764	36.9	61.41	2.98	14.01

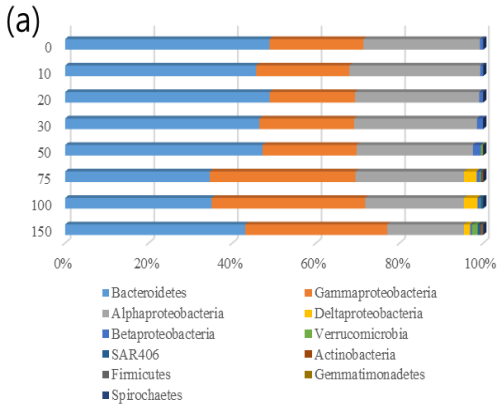
Table 2-11. To be continued

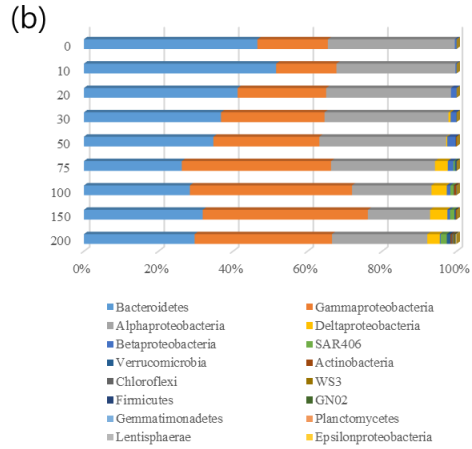
Sample ID	Total reads	No. of OTUs	ACE	Shannon	Simpson
St.9-200m	683	38.7	64.25	3.05	15.05
St.9-300m	319	52.6	88.57	3.35	18.43
St.9-400m	71	nd	nd	nd	nd
St.9-500m	409	50.8	70.37	3.24	14.10
St.9-590m	688	57.8	95.68	3.38	15.16
St.7-0m	902	34.6	47.95	2.82	10.59
St.7-10m	848	33.4	45.22	2.79	10.15
St.7-20m	1165	35.1	49.76	2.91	12.19
St.7-30m	1016	33.3	45.00	2.92	12.94
St.7-50m	1182	35.7	46.80	3.00	13.73
St.7-75m	1432	38.4	48.67	3.14	16.56
St.7-100m	1264	43.7	75.38	3.01	11.29
St.7-150m	1097	41.7	81.05	2.97	11.86
St.7-200m	645	43.3	66.92	3.18	16.11
St.7-300m	995	45.4	70.82	3.23	17.04
St.7-400m	420	44.6	57.90	3.21	15.77
St.7-500m	730	51.4	89.37	3.21	13.72
St.6-0m	1112	37.5	55.54	2.76	9.01
St.6-10m	961	36.8	54.19	2.78	9.56
St.6-20m	1214	43.8	63.02	3.04	10.89
St.6-30m	557	39	49.84	3.10	14.47
St.6-50m	716	43.9	66.93	3.15	14.59
St.6-75m	958	42.7	71.97	3.00	11.01
St.6-100m	761	43.7	77.07	3.05	12.67
St.6-150m	345	41.8	77.76	3.02	13.76
St.6-200m	561	39	61.01	3.03	13.94
St.6-300m	647	44.2	73.75	3.07	12.94
St.6-400m	416	48.9	81.20	3.18	14.01
St.6-500m	355	41.5	69.69	2.92	10.42
St.6-640m	207	58.9	101.48	3.40	15.70
St.5-0m	985	37.5	49.60	2.81	9.03
St.5-10m	1456	42	60.98	3.12	15.12
St.5-20m	982	43.3	62.23	3.14	14.72
St.5-30m	841	43.3	64.80	3.10	13.25
St.5-50m	857	42.5	68.52	3.06	13.00
St.5-75m	887	48.3	98.51	3.09	12.11
St.5-100m	601	41.7	69.93	3.11	15.49
St.5-150m	691	48.4	79.74	3.25	16.68
St.5-200m	508	48.6	74.88	3.24	14.53
St.5-916m	390	53.8	75.63	3.30	14.16

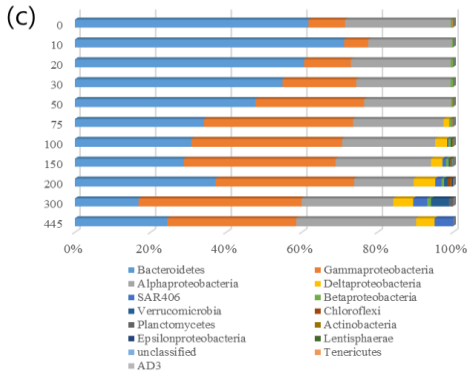
Table 2-12. Pearson correlation coefficient between diversity indices and depth

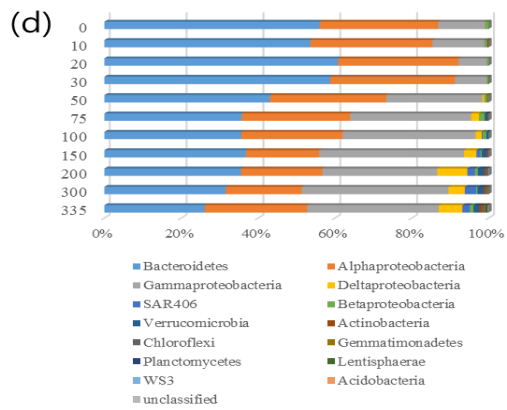
Diversity indices	St.1	St.2	St.3	St.4	St.10	St.8	St.9	St.7	St.6	St.5
No. of OTUs	0.31	0.64	**0.88****	**0.72***	**0.71***	**0.77****	**0.91*****	**0.91*****	**0.74****	**0.77****
ACE	0.35	**0.82****	**0.87****	**0.84****	**0.88****	**0.82****	**0.88*****	**0.70***	**0.76****	0.26
Shannon	0.54	**0.78***	**0.77***	**0.71***	**0.72***	0.60	**0.78****	**0.82****	**0.58***	0.60
Simpson	**0.87****	**0.73***	**0.71***	**0.88****	**0.69***	0.48	0.46	0.53	0.40	0.17

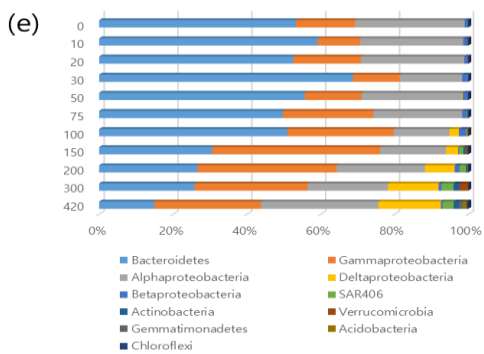
* indicates $p<0.05$, ** $p<0.01$, and *** $p<0.001$, respectively.

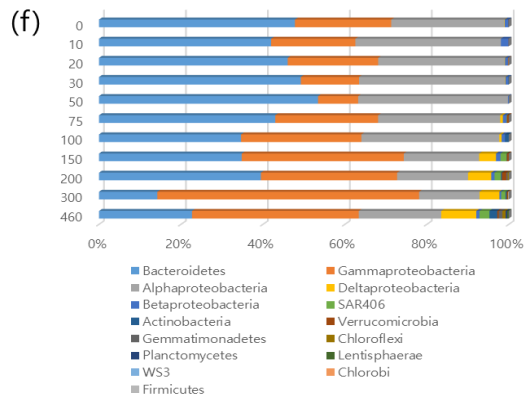

(a)
0
10
20
30
50
75
100
150
0%
20%
40%
60%
80%
100%
Bacteroidetes
Gammaproteobacteria
Alphaproteobacteria
Deltaproteobacteria
Betaproteobacteria
Verrucomicrobia
SAR406
Actinobacteria
Firmicutes
Gemmatimonadetes
Spirochaetes


(b)
0
10
20
30
50
75
100
150
200
0%
20%
40%
60%
80%
100%
Bacteroidetes
Gammaproteobacteria
Alphaproteobacteria
Deltaproteobacteria
Betaproteobacteria
SAR406
Verrucomicrobia
Actinobacteria
Chloroflexi
WS3
Firmicutes
GN02
Gemmatimonadetes
Planctomycetes
Lentisphaerae
Epsilonproteobacteria


(c)
0
10
20
30
50
75
100
150
200
300
445
0%
20%
40%
60%
80%
100%
Bacteroidetes
Gammaproteobacteria
Alphaproteobacteria
Deltaproteobacteria
SAR406
Betaproteobacteria
Verrucomicrobia
Chloroflexi
Planctomycetes
Actinobacteria
Epsilonproteobacteria
Lentisphaerae
unclassified
Tenericutes
AD3


(d)
0
10
20
30
50
75
100
150
200
300
335
0%
20%
40%
60%
80%
100%
Bacteroidetes
Alphaproteobacteria
Gammaproteobacteria
Deltaproteobacteria
SAR406
Betaproteobacteria
Verrucomicrobia
Actinobacteria
Chloroflexi
Gemmatimonadetes
Planctomycetes
Lentisphaerae
WS3
Acidobacteria
unclassified


(e)
0
10
20
30
50
75
100
150
200
300
420
0%
20%
40%
60%
80%
100%
Bacteroidetes
Gammaproteobacteria
Alphaproteobacteria
Deltaproteobacteria
Betaproteobacteria
SAR406
Actinobacteria
Verrucomicrobia
Gemmatimonadetes
Acidobacteria
Chloroflexi


(f)
0
10
20
30
50
75
100
150
200
300
460
0%
20%
40%
60%
80%
100%
Bacteroidetes
Gammaproteobacteria
Alphaproteobacteria
Deltaproteobacteria
Betaproteobacteria
SAR406
Actinobacteria
Verrucomicrobia
Gemmatimonadetes
Chloroflexi
Planctomycetes
Lentisphaerae
WS3
Chlorobi
Firmicutes

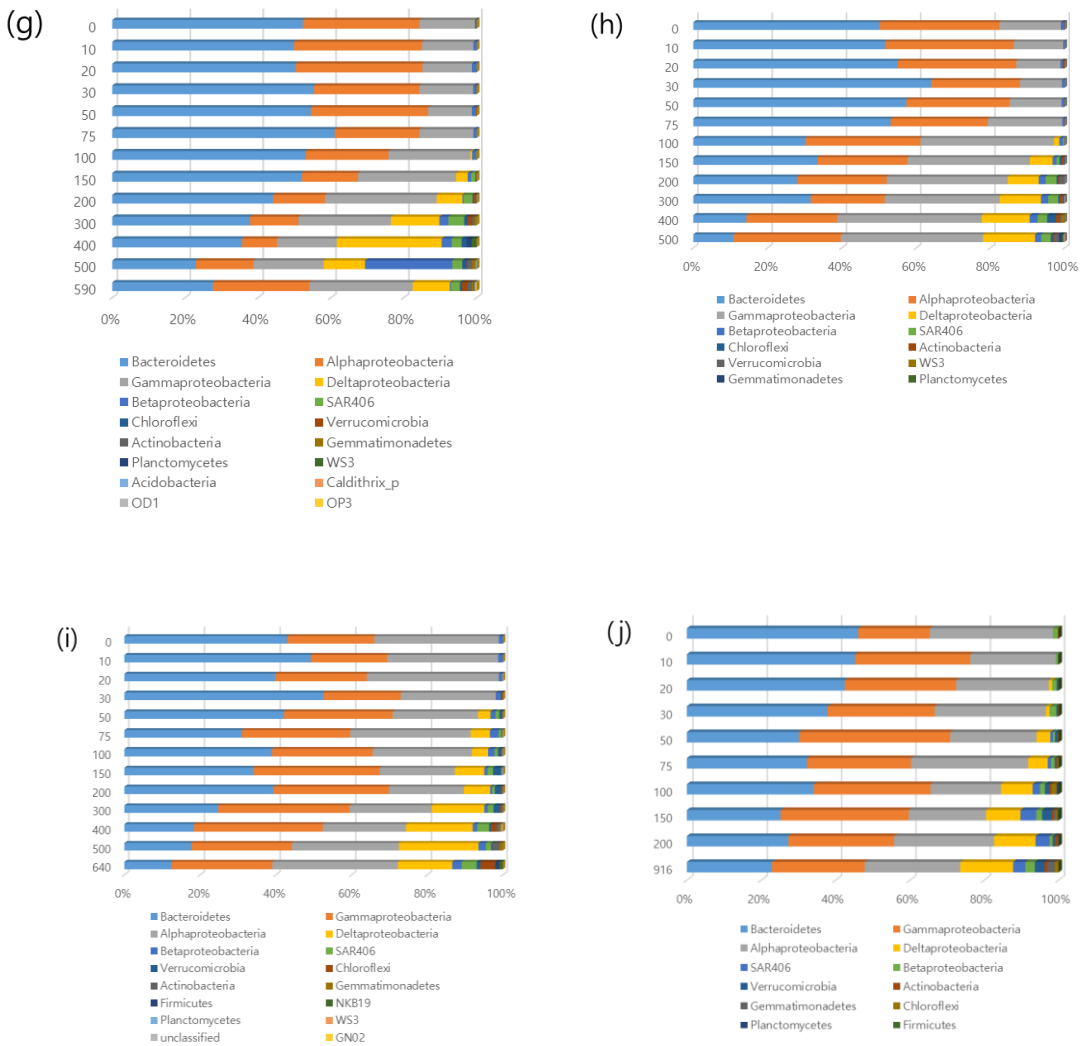


Figure 2-11. Bacterial abundance at the phylum level and class level of *Proteobacteria* at each station (a): Station 1, (b): Station 2, (c): Station 3, (d): Station 4, (e): Station 10, (f): Station 8, (g): Station 9, (h): Station 7, (i): Station 6, (j): Station 5

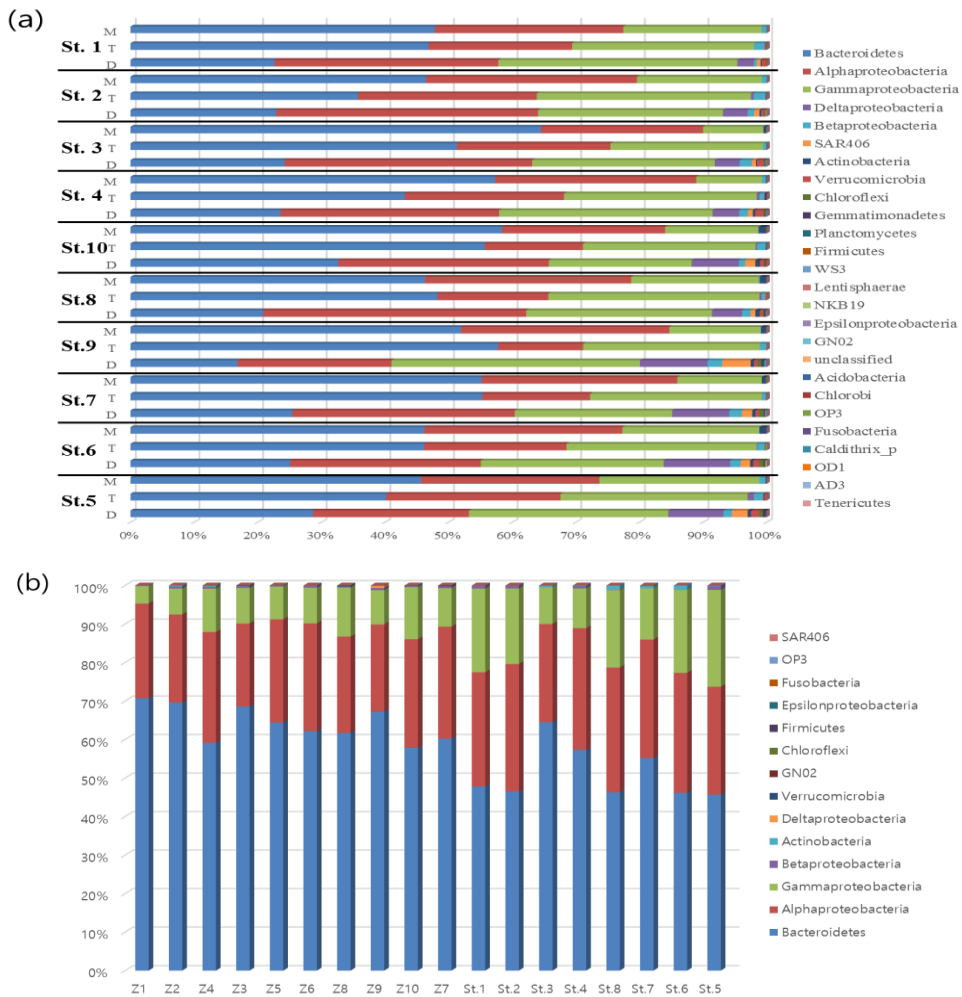


Figure 2-12. Bacterial abundance at the phylum level and class level of *Proteobacteria* according to the water layer divided by environmental variables into mixed layer (M), thermocline (T), and deep water (D) at all stations (a) and average relative abundance in the mixed layer according to the distance from the land (b).

Table 2-13. Relative abundance of phylum or class of *Proteobacteria* according to water layer

Phylum or class of *Proteobacteria*	Mixed layer		Thermocline		Deep water	
	Relative abundance (%)	SD	Relative abundance (%)	SD	Relative abundance (%)	SD
Bacteroidetes	51.98	6.59	48.03	7.14	24.27	4.28
Alphaproteobacteria	30.06	2.81	21.29	5.15	33.78	6.10
Gammaproteobacteria	17.02	5.30	29.09	2.91	30.43	5.16
Deltaproteobacteria	0.02	0.03	0.25	0.31	6.54	3.01
Betaproteobacteria	0.33	0.41	1.09	0.49	1.43	0.54
SAR406	0.00	0.01	0.01	0.02	1.54	1.19
Actinobacteria	0.55	0.47	0.05	0.06	0.45	0.20
Verrucomicrobia	0.00	0.01	0.12	0.12	0.77	0.33
Chloroflexi	0.04	0.05	0.02	0.06	0.35	0.25
Gemmatimonadetes	0.00	0.00	0.03	0.10	0.17	0.16
Planctomycetes	0.00	0.00	0.00	0.00	0.09	0.07
Firmicutes	0.00	0.00	0.00	0.00	0.05	0.07
WS3	0.00	0.00	0.00	0.00	0.05	0.07
Lentisphaerae	0.00	0.00	0.00	0.00	0.02	0.03
NKB19	0.00	0.00	0.01	0.03	0.01	0.03
Epsilonproteobacteria	0.00	0.00	0.00	0.00	0.01	0.02
GN02	0.00	0.00	0.00	0.00	0.01	0.01
unclassified	0.00	0.01	0.00	0.00	0.01	0.01
Acidobacteria	0.00	0.00	0.00	0.00	0.01	0.02
Chlorobi	0.00	0.00	0.00	0.01	0.00	0.01
OP3	0.00	0.01	0.00	0.00	0.00	0.00
Fusobacteria	0.00	0.00	0.00	0.00	0.00	0.01
Caldithrix_p	0.00	0.00	0.00	0.00	0.00	0.01
OD1	0.00	0.00	0.00	0.00	0.00	0.01
AD3	0.00	0.00	0.00	0.00	0.00	0.01
Tenericutes	0.00	0.00	0.00	0.00	0.00	0.00

Table 2-14. Correlation of relative abundance of each phylum or class of *Proteobacteria* with depth at each station

Phylum or class of Proteobacteria	St.1	St.2	St.3	St.4	St.5	St.6	St.7	St.8	St.9	St.10
Bacteroidetes	-0.61	**-0.72***	**-0.79****	**-0.84****	**-0.65***	**-0.92*****	**-0.90*****	**-0.84****	**-0.93*****	**-0.91*****
Alphaproteobacteria	**-0.94*****	**-0.78***	0.25	**-0.73***	-0.09	-0.05	-0.45	**-0.74****	**-0.60***	0.11
Gammaproteobacteria	**0.85****	**0.79***	**0.62***	**0.75****	-0.24	0.41	**0.82****	**0.75****	0.53	0.60
Deltaproteobacteria	0.69	**0.83****	**0.86****	**0.86****	**0.77****	**0.91*****	**0.97*****	**0.95*****	**0.75****	**0.98*****
Betaproteobacteria	-0.35	-0.32	-0.11	-0.05	**0.75***	0.44	**0.85*****	-0.24	0.50	-0.43
SAR406	**0.91****	**0.97*****	**0.98*****	**0.93*****	0.57	**0.88*****	**0.88*****	**0.9*****	**0.80****	**0.96*****
Verrucomicrobia	**0.84***	**0.89****	0.45	**0.91*****	**0.72***	0.29	**0.66***	0.46	**0.66***	0.44
Actinobacteria	0.59	**0.68***	0.04	**0.82****	**0.69***	**0.56***	**0.68***	**0.78****	**0.83*****	**0.91*****
Chloroflexi		**0.76***	0.16	**0.95*****	0.57	**0.74****	**0.69***	**0.86****	**0.81****	**0.70***
Gemmatimonadetes	**0.75***	**0.71***		0.60	**0.98*****	0.53	**0.66***	**0.81****	**0.77*****	**0.73***
Planctomycetes		**0.71***	0.42	**0.88*****	-0.07	0.11	**0.74****	**0.77****	0.39	
WS3		**0.71***		**0.66***		0.30	0.47	0.40	0.32	
Firmicutes	**0.79***	0.31			-0.01	**0.65***		-0.25		
Acidobacteria				**0.62***					**0.74****	**0.73***
Lentisphaerae		**0.71***	0.06	**0.62***			**0.66***	**0.77****		
NKB19						**0.65***				
Epsilonproteobacteria			0.18							
GN02		**0.71***								
unclassified			0.06	0.24		-0.25	**0.66***			
Caldithrix_p									**0.61***	
OD1									**0.61***	
OP3									**0.61***	
Tenericutes			-0.30							
Cyanobacteria							0.28			
AD3			-0.12							
Chlorobi								0.40		

* p<0.05, ** p<0.01, and *** p<0.001, respectively.

The relative abundance of *Bacteroidetes* was mostly represented by class *Flavobacteria* (73.1–100%). The proportions of *Flavobacteria* were 53.5% ± 8.4%, 46.0% ± 8.2%, and 30.0% ± 9.4% in mixed, thermocline, and deep layers, respectively (Table 2-15). *Gammaproteobacteria* was mostly represented by *Oceanospirillales*, *Alteromonadales*, Ruthia_o, and SAR86 orders (58.3–95.2% of *Gammaproteobactetria*). There was no significant change in the portion of *Oceanospirillales* along the depth, whereas the proportions of *Alteromonadales*, Ruthia_o, and SAR86 increased as the depth increased (Table 2-16). The proportion of *Alteromonadales* at the surface was 3.3% and increased to 4.3% and 7.1% in the thermocline and deep waters, respectively. Ruthia_o, which accounted for 1.1% at the surface, increased to 2.7% and 7.8% at thermocline and deeper zones, respectively. Similarly to Ruthia_o, the relative abundance of SAR86 (1.4% at the surface) increased to 1.8% and 5.8% as the depth increased. SAR11 and *Rhodobacterales* accounted for most of the *Alphaproteoacteria* (Table 2-17). The proportion of SAR11 remained constant across the water layers, whereas the relative abundance of *Rhodobacterales* decreased from 11.1% to 4.7% as the depth increased.

Horizontal bacterial community structure

To investigate the horizontal distribution of bacterial composition, surface mixed layers of eight stations (St.1, St.2, St.3, St.4, St.8, St.7, St.6, and St.5) were selected according to the distance from one spot in the land (74.623°S, 164.238°E). Samples from coastal regions (Z1–Z10) were also included for comparisons. The relative abundance of each phylum and class of *Proteobacteria* in surface mixed layers was averaged, and the dynamics according to the distance from the land, which may be an indirect indicator of the input of meltwater runoff resulting from deglaciation, were evaluated. The relative abundance of *Bacteroidetes* decreased as the distance from the coastal line increased ($r = -0.65$, $p < 0.01$), and the portions of *Gammaproteobacteria* ($r = 0.69$, $p < 0.01$) and

Table 2-15. Relative abundance of classes of *Bacteroidetes* according to the water layer

Order	Mixed layer		Thermocline		Deep layer	
	MRA†	SD*	MRA†	SD*	MRA†	SD*
Flavobacteria	53.5	8.4	46	8.2	30	9.4
Cytophagia	1	0.5	1.8	0.8	0.9	0.7
Sphingobacteria	0.1	0.1	0.1	0.1	0.1	0.1
Bacteroidia	0	0	0	0	0.1	0.7

† MRA represents mean relative abundance.

* SD means the standard deviation among mean relative abundance of each station.

Table 2-16. Relative abundance of orders of *Gammaproteobacteria* according to the water layer

Order	Mixed layer		Thermocline		Deep layer	
	MRA†	SD*	MRA†	SD*	MRA†	SD*
Oceanospirillales	8.7	2.8	9.1	2.5	7.6	3
Alteromonadales	3.3	1.6	4.3	1.3	7.1	6.4
Ruthia_o	1.1	1	2.7	2.2	7.8	2.9
SAR86	1.4	1.2	1.8	1.5	5.8	2.3
OM182_o	0.6	0.4	1.6	0.8	2.3	1.2
Thiotrichales	0.4	0.4	0.4	0.4	0.9	1.1
OM60_o	0.3	0.2	0.7	0.5	0.4	0.5
DQ906757_o	0.5	0.3	0.7	0.3	0.1	0.2
Steroidobacter_o	0	0	0	0	0.6	0.8
Chromatiales	0	0.1	0.1	0.2	0.3	0.4
EF572656_o	0	0	0.1	0.1	0.3	0.2
unclassified	0	0	0	0.1	0.1	0.2
HQ671930_o	0	0	0	0	0.1	0.2
Pseudomonadales	0	0.1	0.1	0.1	0	0.1
EF573277_o	0	0	0	0	0.1	0.2
Legionellales	0	0	0	0	0	0.2
Arenicella_o	0	0	0	0	0	0.1

† MRA represents mean relative abundance.

* SD means the standard deviation among mean relative abundance of each station.

Table 2-17. Relative abundance of orders of *Alphaproteobacteria* according to the water layer

Order	Mixed layer		Thermocline		Deep layer	
	MRA†	SD*	MRA†	SD*	MRA†	SD*
SAR11	18.79	4.45	18.4	4.57	16.51	4.36
Rhodobacterales	11.05	1.94	9.7	2.13	4.71	2.87
Rhodospirillales	0.15	0.2	0.32	0.3	1.03	1.14
SAR116	0.18	0.17	0.36	0.31	0.25	0.4
Kordiimonadales	0.01	0.04	0.1	0.15	0.17	0.21
Rhizobiales	0	0	0.02	0.06	0.14	0.24
EF573150_o	0	0	0.01	0.04	0.05	0.1
Rickettsiales	0	0	0	0	0.02	0.07
Sphingomonadale	0	0	0	0	0.01	0.06

† MRA represents mean relative abundance.

* SD means the standard deviation among mean relative abundance of each station.

Actinobacteria (r = 0.60, $p < 0.01$) increased. The highest proportion was *Bacteroidetes* (59.9–70.7%), and the lowest proportion was *Gammaproteobacteria* (4.6–13.5%) in the coastal areas of TNB (Z1–Z10; Figs. 2-11 and 2-12b). The proportion of *Alphaproteobacteria* was not significantly changed and ranged from 21.5% to 33.1% along the distance.

Bacterial community structure and relation to environmental variables

ANOSIM analysis revealed that the bacterial community structures of each layer could be distinguished (ANOSIM Global $R = 0.6$, $P < 0.001$). Analysis of the relationship between bacterial composition and environmental variables, such as inorganic compound concentration (phosphate, nitrite, nitrate, silicate, and ammonium), temperature, salinity, fluorescence, oxygen concentration, and pressure revealed that temperature, oxygen, and fluorescence, positively influenced the bacterial community structure of the mixed layer, whereas salinity and phosphate, nitrite, nitrate, and silicate concentrations affected the bacterial community structure of the deep layer (Fig. 2-13). In particular, oxygen concentration explained the highest variance in the bacterial community structure ($r^2 = 0.8$, $P < 0.0001$), followed by the concentrations of nitrite and nitrate ($r^2 = 0.7$, $P < 0.0001$) along the water depth. Notably, the temperature and salinity of the surface water decreased as the distance from the land increased, whereas the concentrations of inorganic nutrients, such as phosphate, nitrate, nitrite, and silicate increased (Table 2-1 and Fig. 2-14), supporting the changes in bacterial compositions of *Bacteroidetes*, *Gammaproteobacteria*, and *Actinobacteria* (Table 2-18).

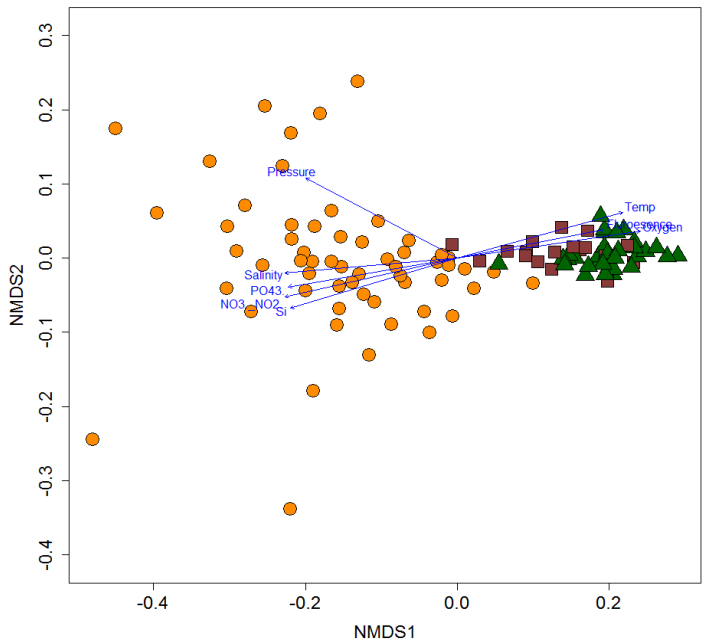


Figure 2-13. NMDS analysis of bacterial composition and PCA analysis. Environmental variables used are concentration of inorganic compounds (phosphate, nitrite and nitrate, silicate, and ammonium), temperature, salinity, fluorescence, concentration of oxygen, and pressure. Green triangles indicate samples from mixed layers, brown squares for thermocline, and orange circles are for deep layers.

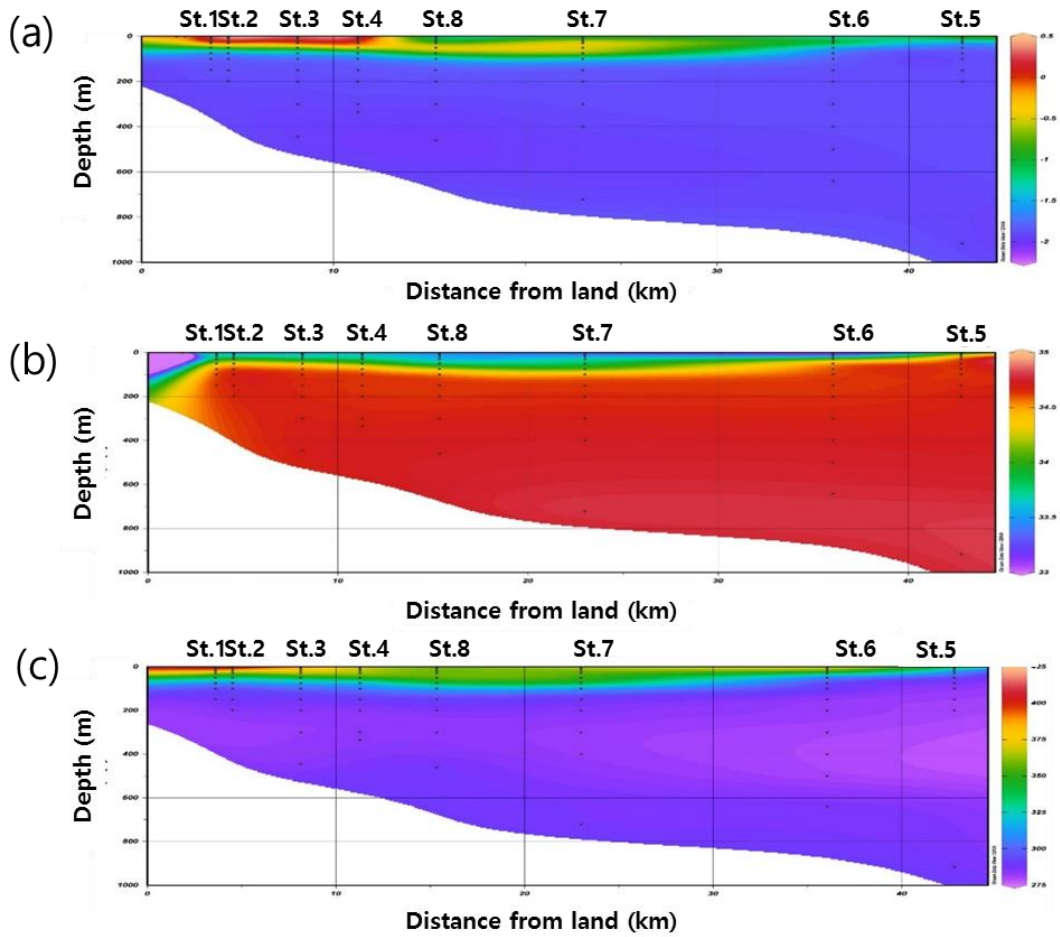


Figure 2-14. Spatial profiles of potential temperature (a), salinity (b), and oxygen (c). Sampling depths in each site were identified by closed circles. The contours were drawn from data collected from the cruise R/V ARAON in 2011.

Table 2-18. Correlation of relative abundance of each phylum or class of *Proteobacteria* with the distance from the land

Phylum or class of Proteobacteria	Pearson correlation coefficient	*p*
Bacteroidetes	**-0.65**	**0.0037**
Alphaproteobacteria	0.40	0.1030
Gammaproteobacteria	**0.69**	**0.0016**
Betaproteobacteria	-0.01	0.9593
Actinobacteria	**0.60**	**0.0091**
Deltaproteobacteria	-0.13	0.6099
Verrucomicrobia	-0.09	0.7261
GN02	-0.26	0.2982
Chloroflexi	0.30	0.2279
Firmicutes	-0.14	0.5926
Epsilonproteobacteria	-0.20	0.4277
Fusobacteria	-0.14	0.5679
OP3	-0.02	0.9432

2.3.3 Vertical and spatial distribution of bacterial composition in marine sediments of the Ross Sea

Environmental factors along the depth

The contents of water, TOC, and TN of the core varied according to the lithologic characterisitcs of sediment. The core can be divided into three units: the upper unit of greenish gray diatomaceous mud with minor ice-rafted debris (core depth: 0–120 cmbsf), the middle unit of dark greenish gray to dark gray diamicton and sandy mud (120–360 cmbsf), and the lower unit of light greenish gray diatomaceous mud (360–396 cmbsf). Clay- and silt-sized grains were dominant throughout the core, but the content of coarser-sized grains (sands and gravels) reached up to 20% in the middle unit. The upper unit showed typical characteristics of postglacial sediment facies in the western Ross Sea indicating a deposition under a seasonally open-marine conditions while the middle and lower units represent sedimentation during the last glacial and deglacial periods under a glacial-proximal setting with a short period of seasonally open-marine conditions. Water content of upper unit sediment (63-64 wt.%) was much higher than that of middle and lower unit sediments (~30 wt.%) (Fig. 2-15), which can be attributed to differences in grain size distribution and depositional setting. TOC content was less than 1 wt.% throughout the core (Fig. 2-15). Contents of TOC and TN were higher in the upper unit than in the lower units (Fig. 2-15), which may be related with higher primary production during postglacial periods than glacial periods.

Patterns of phyla and major OTUs distribution

Taxonomic assignment revealed the presence of 94 distinct bacterial lineages and 68 out of them were not assigned to any known phylum indicating they are candidate

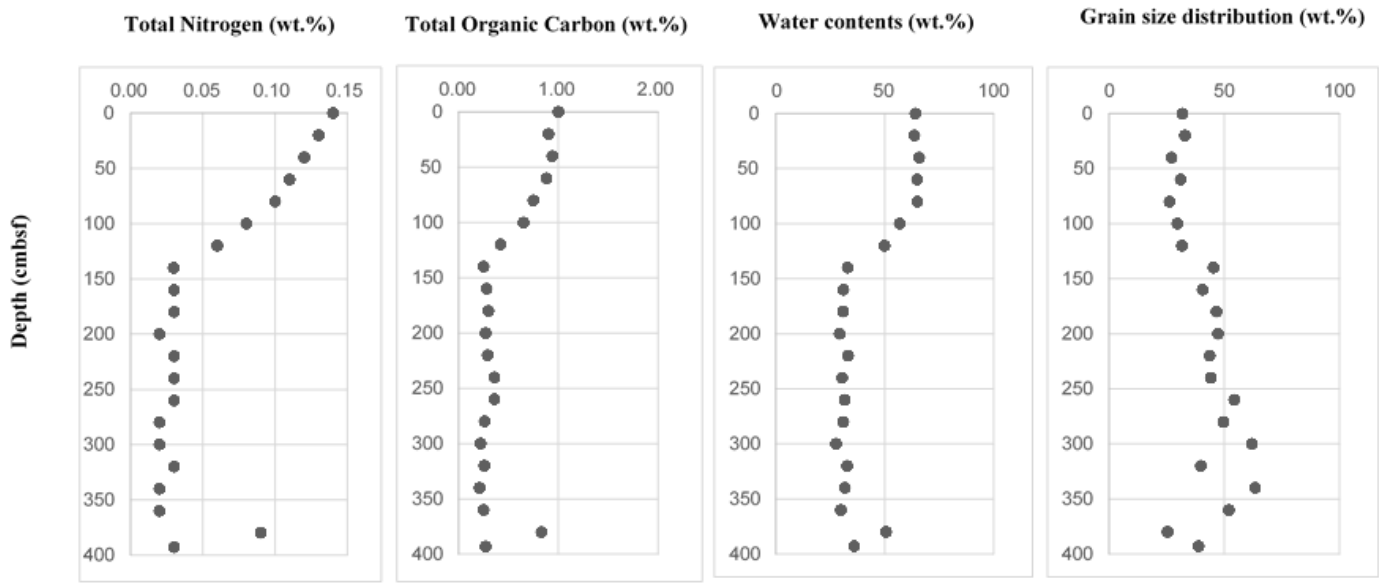


Figure 2-15. Depth profile of the concentration of total nitrogen, total organic carbon, water contents and grain size distribution. Grain size distribution was measured by the content of silt- and clay-sized part.

phylum with no cultured representatives (Fig. 2-16a). *Gammaproteobacteria*, candidate division JS1, *Chloroflexi*, *Actinobacteria*, *Deltaproteobacteria*, *Planctomycetes*, *Alphaproteobacteria*, *Bacteroidetes*, and *Betaproteobacteria* were the predominant groups, comprising 63.9-91.9% of the bacterial communities (Fig. 2-16a). There were great differences in the dominant phyla between aerobic and anoxic environments (Fig. 2-16b). In aerobic sediments, *Gammaproteobacteria*, *Deltaproteobacteria*, *Bacteroidetes*, *Planctomycetes*, and *Alphaproteobacteria* were dominant (45.6%-85.6%) while the relative abundance of those phyla ranged from 7.0% to 29.7% in anaerobic environments. By contrast, dominant groups in anoxic sediments were candidate division JS1, *Chloroflexi*, *Actinobacteria*, *Betaproteobacteria*, and *Alphaproteobacteria* (5.8%-35.2%) and the portion of candidate division were higher (21.4%-56.2%) than that of surface samples (1.3-20.3%) (Table 2-19).

When the bacterial community was analyzed along the depth within a core, a distinctive phyla stratification along the depth was revealed (Fig. 2-16a and Fig. 2-17). *Proteobacteria*, *Planctomycetes*, *Bacteroidetes*, *Acidobacteria*, and *Chlorobi* were more abundant in the oxic upper sediment, whereas Ca. *Atribacteria* JS1, *Chloroflexi*, and *Actinobacteria* were found in higher abundance in the anoxic sediment horizon. *Alphaproteobacteria* (21.2%), *Deltaproteobacteria* (12.9%), *Gammaproteobacteria* (11.2%), *Planctomycetes* (8.9%), *Bacteroidetes* (8.3%), and *Actinobacteria* (7.4%) were the six most predominant groups on the surface, whereas Ca. *Atribacteria* JS1 (9.0–39.6%), *Chloroflexi* (5.8–26.0%), *Actinobacteria* (1.9–33.7%), and candidate division OP8 (0.8–18.1%) were abundant groups in sediments at and below 20 cmbsf (Fig. 2-16a and 2-17a). Candidate phyla accounted for 9.0% of the bacterial community at the surface and 21.4–56.2% in anoxic sediments (≥ 20 cmbsf). Among candidate phyla, Ca. *Atribacteria* JS1, OP8, and CD12 accounted for the most of the candidate phyla (17.2–46.9%). Ca. *Atribacteria* JS1 was rarely recovered at 0 cmbsf (0.02%), but ranged from 9.0% to 39.6% at and below 20 cmbsf, showing a clear niche separation within the core.

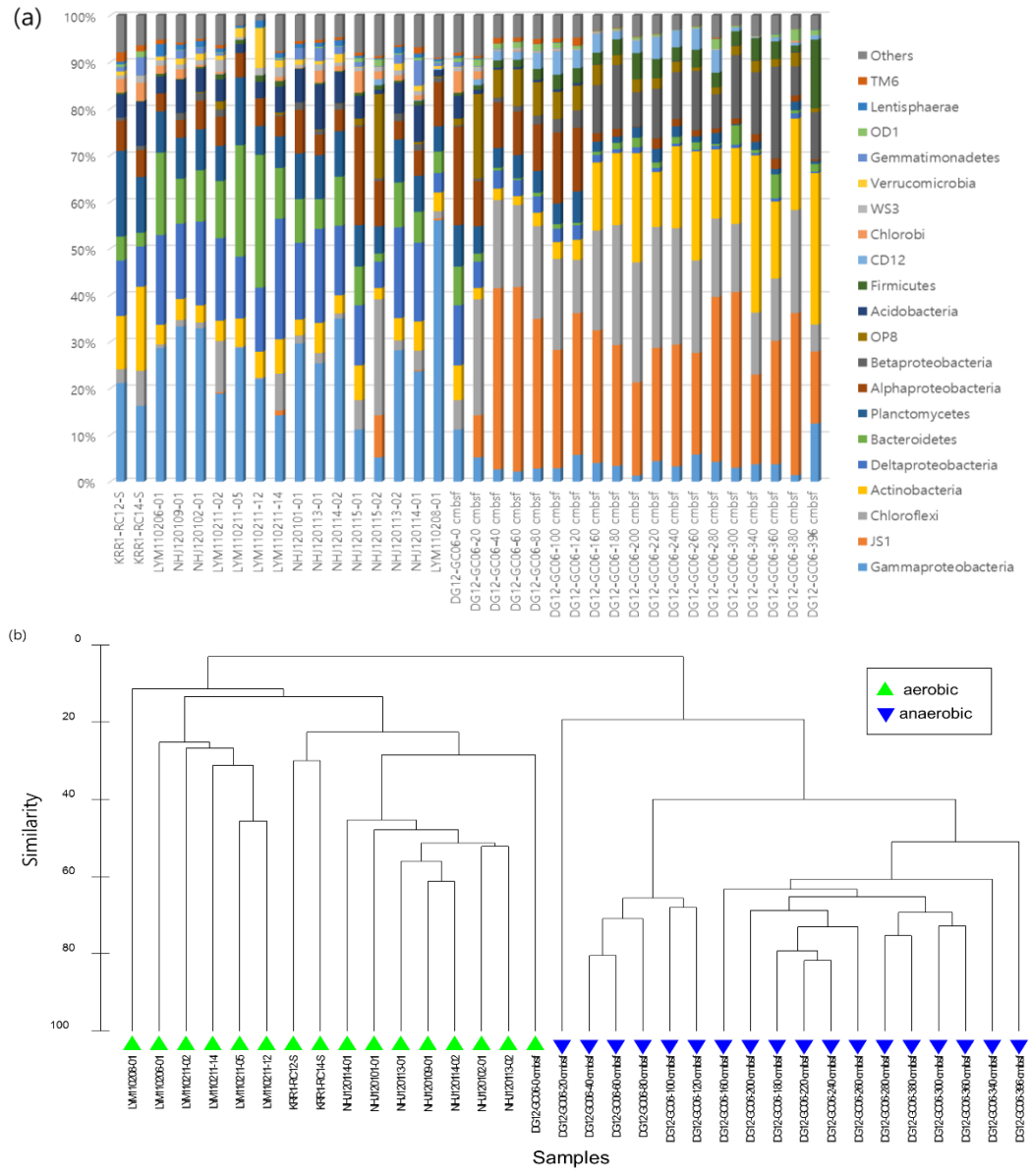


Figure 2-16. Bacterial communities represented by the relative abundance at the phylum level and class level of *Proteobacteria* (a) and similarities of samples by bacterial composition based on relative abundance of each OTUs (b)

Table 2-19. Relative abundance according to the presence of oxygen

Phylum and classes of Proteobacteria	Aerobic		Anaerobic	
	Average	SD	Average	SD
Gammaproteobacteria	25.14	11.75	4.00	2.47
JS1	0.12	0.27	27.28	8.36
Chloroflexi	3.22	3.11	18.99	5.68
Actinobacteria	6.07	3.77	14.29	10.03
Deltaproteobacteria	17.10	7.87	1.78	1.56
Bacteroidetes	10.70	6.87	1.44	1.29
Planctomycetes	9.52	3.11	2.84	1.87
Alphaproteobacteria	6.33	4.27	4.68	5.02
Betaproteobacteria	0.87	0.61	8.14	5.47
OP8	0.54	1.07	4.23	4.09
Acidobacteria	5.44	2.45	0.22	0.27
Firmicutes	0.73	0.78	3.67	3.01
CD12	0.03	0.08	2.60	1.67
Chlorobi	1.46	1.39	0.16	0.39
WS3	1.26	0.89	0.23	0.22
Verrucomicrobia	1.43	1.92	0.01	0.02
Gemmatimonadetes	1.50	1.49	0.03	0.05
OD1	0.34	0.33	0.84	0.70
Lentisphaerae	1.07	0.67	0.04	0.06
TM6	0.69	0.49	0.39	0.54
Others	6.42	3.20	4.13	1.48

Twenty one major OTUs with 3% or greater portion in at least two samples were selected (Fig. 2-18). In aerobic samples, relative abundance of each major OTU did not exceed 13.5% and three out of six major OTUs were recovered in samples from Australian-Antarctic Ridge and they were belonged to *Actinobacteria*, *Chlorobi*, and *Gammaproteobacteria* while other three major OTUs from the Ross Sea belonged to *Alphaproteobacteria*, *Gammaproteobacteria* and *Deltaproteobacteria* (Fig. 2-18). With the exception of one alphaproteobacterial OTU, c11696 and gammaproteobacterial OTU, c1477 which were mostly recovered in the oxic sediments, the distribution of major OTUs in oxic sediments from the Ross Sea showed different distribution pattern according to the location. Major OTUs from anaerobic environments accounted for up to 38.6% and some major OTUs showed clear niche separation within a core. c4865 of candidate division JS1 that was rarely recovered from aerobic samples, dominated throughout the core composing up to 38.6% while c7832 of *Actinobacteria* represented from 10.0% to 30.9% from 140 cmbsf to 400 cmbsf. Similarly with c7832, c3393 and c558 belonged to the *Burkholderiales* of *Betaproteobacteria* were predominant at the deeper part of the core. Most of the OTUs of *Actinobacteria*, *Bacteroidetes*, *Chlorobi*, *Chloroflexi*, *Firmicutes*, *Alphaproteobacteria*, *Betaproteobacteria*, *Gammaproteobacteria*, and candidate divisions JS1, CD12, and OP8 showed higher portion in anoxic environments. The number of OTUs recovered in the aerobic environments was 8,875 and that from the anaerobic samples were 3,359. The number of shared OTUs between two environments was 1,012 (9% of total OTUs recovered) (Fig. 2-19) indicating these environments harbor the distinct OTUs.

OTU_ID	phylum	class	order	family	genus	species	Aerobic																Anaerobic																						
							KRR1-RC12-S	KRR1-RC14-S	LYM11 0206-01	NHJ12 0109-01	NHJ12 0102-01	LYM11 0211-02	LYM11 0211-05	LYM11 0211-12	LYM11 0211-14	NHJ12 0101-01	NHJ12 0113-01	NHJ12 0114-02	NHJ12 0114-03	NHJ12 0113-02	NHJ12 0114-01	LYM11 0208-01	DG12-GC06-0 cmbsf	DG12-GC06-20 cmbsf	DG12-GC06-40 cmbsf	DG12-GC06-60 cmbsf	DG12-GC06-80 cmbsf	DG12-GC06-100 cmbsf	DG12-GC06-120 cmbsf	DG12-GC06-160 cmbsf	DG12-GC06-180 cmbsf	DG12-GC06-200 cmbsf	DG12-GC06-220 cmbsf	DG12-GC06-240 cmbsf	DG12-GC06-260 cmbsf	DG12-GC06-280 cmbsf	DG12-GC06-300 cmbsf	DG12-GC06-340 cmbsf	DG12-GC06-360 cmbsf	DG12-GC06-380 cmbsf	DG12-GC06-396 cmbsf
c7832	Actinobacteria	OPB41	EU290722_o	EU290722_f	EU290722_g	EU385678_s	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.26	10.95	13.07	21.17	9.97	15.44	21.13	13.52	12.50	30.93	12.64	18.22	21.92
c13635	Actinobacteria	EU374107_c	EU374107_o	EU374107_f	EU374107_g	GQ850565_s	4.38	3.53	0.00	1.25	1.64	1.27	2.20	0.63	2.93	1.13	0.65	0.79	0.04	1.02	0.45	0.28	1.22	0.11	0.02	0.07	0.02	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
c2807	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	Cloacibacterium	Cloacibacterium_normanense	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.14	0.11	0.15	0.32	0.61	0.23	0.16	0.94	1.85	0.98	0.91	1.04	0.29	3.71	0.27	4.77	0.38	1.10
c11228	CD12	CD12_c	CD12_o	ASMN_f	ASMN_g	ASMN_s	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.11	0.00	0.00	1.16	1.46	1.68	3.22	4.85	3.48	3.35	1.51	2.11	3.84	3.08	3.85	3.70	0.66	0.17	0.44	0.48	0.67
c1203	Chlorobi	Ignavibacteriae	Ignavibacteriales	Ignavibacteriaceae	Ignavibacterium	unclassified	2.86	3.38	0.00	1.96	0.28	0.45	0.00	0.00	0.30	0.12	2.50	1.02	0.94	0.85	0.92	0.00	4.27	1.28	0.04	0.18	0.07	0.04	0.00	0.00	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
c3064	Chloroflexi	Dehalococcoidetes	Dehalococcoidales	GQ259278_f	DQ394968_g	4P000437_s	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.24	0.14	0.00	0.01	0.07	1.65	9.00	8.18	11.07	9.97	5.17	1.28	0.12	0.60	0.06	0.90	0.51
c2483	Chloroflexi	Dehalococcoidetes	Dehalococcoidales	GQ259278_f	DQ394968_g	AY093456_s	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	1.89	1.60	0.56	1.21	0.02	0.04	1.27	1.90	5.07	2.93	4.49	4.40	6.89	4.63	2.16	4.25	0.67
c7656	Chloroflexi	Anaerolineae	Anaerolinaeles	AM745150_f	AM745150_g	AM745150_s	0.00	0.00	0.00	0.00	0.00	0.27	0.00	0.00	0.98	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.17	4.81	4.90	4.72	2.25	5.13	3.71	0.08	0.00	0.05	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00
c5189	Chloroflexi	Dehalococcoidetes	Dehalococcoidales	GQ259278_f	4P000399_g	4P000399_s	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.04	0.11	3.05	1.00	0.90	5.45	1.31	0.69	0.44	0.44	1.13	0.52	0.00	0.00	0.55	0.23	0.16
c4865	JS1	JS1_c	JS1_o	JS1_f	JS1_g	unclassified	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.78	37.79	38.57	31.28	24.83	29.68	27.01	25.60	19.53	23.92	25.52	20.59	35.15	37.46	18.65	26.27	33.76	15.18
c14594	JS1	JS1_c	JS1_o	JS1_f	JS1_g	unclassified	0.00	0.00	0.00	0.01	0.01	0.27	0.00	0.00	1.05	0.00	0.00	0.00	0.01	0.00	0.25	0.14	0.02	6.16	1.03	0.88	0.81	0.57	0.72	1.32	0.26	0.44	0.37	0.60	1.22	0.23	0.25	0.56	0.22	0.97	0.28
c10521	OP8	OP8_c	OP8_o	OP8_f	AQSP_g	4P004969_s	0.00	0.00	0.00	0.03	0.00	0.09	0.00	0.00	0.08	0.00	0.20	0.00	0.03	0.00	0.22	0.14	0.15	9.37	2.38	2.21	4.46	2.22	3.05	3.36	1.42	2.41	1.67	2.05	0.59	1.24	1.57	2.07	1.50	1.68	0.83
c1655	OP8	OP8_c	OP8_o	OP8_f	AQSP_g	EU735011_s	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.07	0.00	0.02	0.00	0.73	4.79	1.48	2.87	0.95	1.97	0.50	0.01	0.04	0.05	0.02	0.00	0.00	0.05	0.00	0.00	0.00	1.07	0.00
c11696	Proteobacteria	Alphaproteobacteria	Rhizobiales	Hyphomicrobiaceae	Methyloligella	GQ357015_s	0.67	1.13	1.54	0.73	1.29	4.63	2.75	0.76	2.63	1.41	1.48	1.41	0.88	0.60	2.11	1.95	13.48	3.31	4.13	4.36	4.97	7.70	6.29	0.05	0.02	0.01	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.08
c3393	Proteobacteria	Betaproteobacteria	Burkholderiales	Ralstonia_f	Ralstonia	AM713401_s	0.00	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.21	0.43	0.43	0.62	0.97	0.89	4.85	6.37	0.27	4.14	3.78	3.17	3.41	3.71	3.94	5.71	2.68	0.28
c558	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Comamonas	Comamonas_denitrificans	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.00	0.17	0.28	0.21	0.70	0.59	0.76	3.96	5.82	3.40	3.50	6.35	1.51	3.67	0.81	6.60	1.80	4.97
c6469	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	Halomonas_desiderata	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.02	0.10	0.00	0.39	0.05	0.13	0.00	0.10	0.34	0.92	1.20	1.30	1.31	2.42	2.54	2.04	0.60	2.62	1.97	3.58	2.16	1.57	2.20	1.94	0.78	8.55
c1477	Proteobacteria	Gammaproteobacteria	Steroidobacter_o	AB013829_f	AB013829_g	AB013829_s	3.14	5.71	0.00	1.97	1.10	0.36	0.27	1.14	2.18	1.39	2.11	1.75	0.03	1.28	1.70	0.97	1.07	0.07	0.02	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
c3764	Proteobacteria	Gammaproteobacteria	Steroidobacter_o	AB013829_f	AB013829_g	unclassified	0.00	0.53	0.00	3.67	2.20	0.00	0.00	0.00	0.00	1.87	3.24	4.40	0.09	2.10	4.42	0.00	2.93	0.09	0.02	0.01	0.03	0.00	0.05	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
c1371	Proteobacteria	Alphaproteobacteria	Rhizobiales	Hyphomicrobiaceae	Hyphomicrobium	FJ197410_s	2.19	0.90	0.00	0.14	0.10	0.18	0.00	0.00	0.00	0.07	0.13	0.15	0.12	0.03	0.15	0.56	1.44	2.67	1.89	1.97	1.95	3.84	4.29	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.08
c8581	Proteobacteria	Deltaproteobacteria	Desulfobacterales	Desulfobacteraceae	AJ704681_g	AJ704681_s	0.48	0.68	0.00	1.44	0.43	0.73	0.55	0.38	3.38	0.05	2.09	0.66	5.06	0.05	1.02	0.28	1.95	0.78	0.60	0.43	0.27	0.43	0.20	0.18	0.02	0.07	0.15	0.10	0.00	0.13	0.04	0.04	0.00	0.05	0.04

Figure 2-17. Heat map of major OTUs. Major OTUs with 3% or greater compositions at least in two samples were selected.

The color gradient from white to brown indicates the lowest to highest relative abundance values.

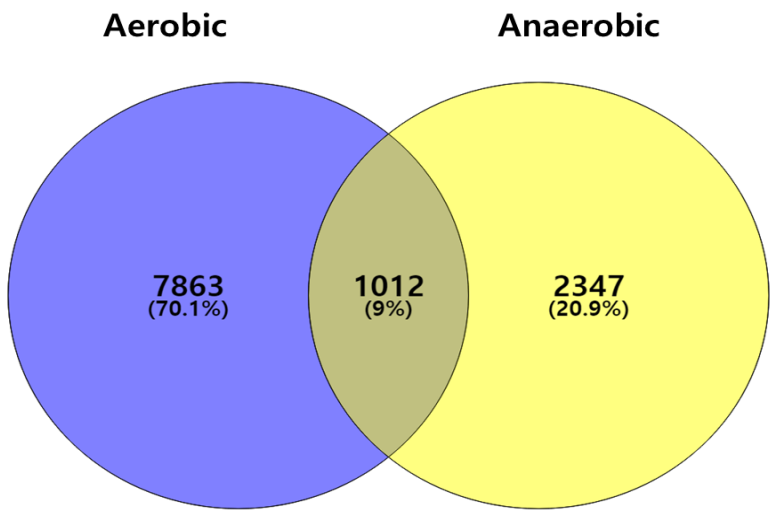


Figure 2-18. Venn diagram showing the overlap in bacterial OTUs between aerobic and anaerobic sediments

2.4 Discussion

2.4.1 Biogeographic structure of the bacterial communities according to the circumpolar fronts in the Southern Ocean

In this study aimed to investigate whether the bacterial communities of the surface seawater vary along the fronts in the SO, bacterial communities were significantly separated by the SAF and SACCF and the proportions of some bacterial phyla were strongly affected by the frontal zones. Bacterial community assemblages constructed by environmental filtering could be supported by differentiating samples into 2 distinct clusters in the SAF, based on environmental variables in transect 1 and the SACCF in transect 2, which matched well with the separation of the microbial community. The SAF is the boundary between the sub-Antarctic and Antarctic and the northernmost front of the Antarctic Circumpolar Current (ACC). The ACC is the largest and strongest current system in the world oceans and serves as a conduit of active and passive tracers linking the ocean basins of the Atlantic, Pacific, and Indian Oceans and isolates Antarctica from the other oceans (Sokolov and Rintoul 2007).

The relative abundance of *Bacteroidetes* increased while that of *Cyanobacteria*, *Verrucomicrobia*, *Actinobacteria*, and SAR406 decreased as the latitude increased. Regardless of the proportion, the relative abundance of *Cyanobacteria*, *Verrucomicrobia*, *Actinobacteria*, or SAR406 shifted dramatically by presence of the fronts indicating they were niche-specifically adapted more than other groups. The high abundance of *Bacteroidetes*, specifically members of the class *Flavobacteria* (95.6 – 100%) throughout the transects, and no clear changing pattern across each front despite the overall increase along the latitudinal increase explained why members of this phylum may be habitat generalists that respond predominantly to environmental gradients (Logares et al. 2013). The higher proportion of *Flavobacteria*, which are heterotrophic degraders of high molecular-weight compounds seemed to be strongly linked to high nutrient concentration resulting from algal blooms in the colder seawaters and indicated their important role in

nutrient cycling in marine ecosystems.

Cyanobacteria, which are known to represent a substantial fraction of primary marine production and carbon cycling, were represented by a single dominant OTU of the genus *Prochlorococcus* and were not recovered from seawaters south of the SAF in this study. The distribution pattern of *Prochlorococcus* observed in this study was consistent with that of previous studies in that *Prochlorococcus* typically dominated in warm oligotrophic waters and the regional distribution was mainly controlled by temperature, which contributed significantly to the primary productivity of the tropical and subtropical oceans (Flombaum et al. 2013, Huang et al. 2012, Partensky et al. 1999, Thomas et al. 2012). As proposed by Wilkins et al. (2012), the primary production of the small-sized bacteria *Prochlorococcus* in lower latitudes may be replaced by larger phytoplankton with higher intrinsic maximum growth rates in higher latitudes, which was advantageous for their production (Follows et al. 2007). This possibility is partially supported by previous data showing a direct correlation for prokaryotic autotroph growth with temperature in the SO, but a high abundance of eukaryotic cells in the coldest high-latitude waters, although it is not yet clear which factor provides a competitive advantage to the eukaryotic picoplankton over prokaryotes (Marchant and McEldowney 1986).

Conversely to the result of Wilkins et al. (2012), in the current study, *Verrucomicrobia* was more abundant in the north of PF and showed a clear biogeographic pattern being observed mostly only below 57°S, approximate location of PF. This indicates PF functions as a biogeographic barrier to members of *Verrucomicrobia* in the SO. Previously, in spite of the ubiquity of *Verrucomicrobia* across a range of environmental conditions, lineages abundant in the water column was positively correlated with the temperature in agreement with our results (Freitas et al. 2012). Since little is known about the ecophysiology of *Verrucomicrobia* lineages, reasons about the dominance of one verrucomicrobial OTU in warmer seawaters and clear biogeographical distribution fashions according to frontal zones in both transects of this study in link with the ecological

roles need to be elucidated.

This study was performed to give a clear answer about the biogeographic barrier for bacterial community structure in a wide range of the latitude of the SO by sampling across the fronts at higher resolution. SAF and SACCF, the northernmost and southernmost boundary of Antarctic Circumpolar Current boundary played as the strong biogeographic barrier to bacterial communities in SO despite of the longitudinal variance. The taxonomic distribution pattern by the frontal systems indicates different and distinct metabolic traits and functions exist in each frontal zone and thus, the climate change would drive the shifts in bacterial communities resulting in ecological function in SO that may influence at the global scale.

2.4.2 Vertical and horizontal distribution of bacterial communities in seawater of the Ross Sea

In this study, we observed high abundances of *Bacteroidetes*, *Alphaproteobacteria*, and *Gammaproteobacteria* dominated throughout the depth at all stations, consistent with the results from previous studies in Antarctic waters, including the Ross Sea (Abell and Bowman 2005, Celussi et al. 2009a, Gentile et al. 2006, Ghiglione et al. 2012, Murray and Grzymski 2007). In addition to the dominance of these groups, a clear stratification in bacterial community compositions with an increase of diversity along the water column was observed. A reduction in the total proportions of the three dominant groups, i.e., *Bacteroidetes*, *Alphaproteobacteria*, and *Gammaproteobacteria*, along the depth was followed by an increase in the proportions of *Betaproteobacteria*, SAR406, *Actinobacteria*, *Verrucomicrobia*, *Chloroflexi*, *Gemmatimonadetes*, *Planctomycetes*, and *Firmicutes*, indicating that the metabolic traits shifted according to the physicochemical factors.

Bacterial community clustering mirrored the changes in environmental parameters (i.e., oxygen concentration and nitrite/nitrate concentrations) as the depth increased. Surface water, which exhibits high concentrations of Chl a, an indirect indicator of primary production, and high oxygen levels compared with those of deep water, may provide optimal growth opportunities to heterotrophs using high-molecular-weight molecules and aerobes; in contrast, deep waters with low levels of oxygen and high concentration of inorganic nutrients may form a niche that is preferable to facultative anaerobes with versatile metabolic traits using diverse substrates for energy conservation, resulting in a rich diversity in the deep waters of the ocean. Consistent with our results, the lower divergence in the 16S rRNA genes in surface water compared with that of deep water has been reported in previous works, indicating the presence of a specialized bacterial population with a higher tolerance to dynamic environmental changes that may adapt to surface water (Celussi et al. 2009a, Ghiglione et al. 2012).

Compared with the clear vertical stratification in bacterial communities within water columns, the horizontal distribution of the bacterial composition did not differ as much. Although there were variations in physical variables in the surface layers among stations, this variation was less than that within water columns at each station, supporting the high similarities among bacterial communities of surface seawaters. Since the parameters for terrestrial input and fluorescence near costal line samples (Z1–Z10) were not measured in this study, our interpretation of the horizontal distribution depended only on the distance measured in a one-way direction from the coastal line and parameters excluding primary production. However, because there was no sea ice at the sampling sites during the sampling period, the high portion of *Bacteroidetes* at near the coast, known for its strong association with the algal bloom, may imply the occurrence of an algal bloom in TNB, resulting in horizontal differences in *Bacteroidetes* as the distance from the land changed.

The portions of *Bacteroidetes* decreased as the water depth changed from the

mixed layer to deep water. The majority of *Bacteroidetes* were represented by members of the class *Flavobacteria* across all depths. The high portions of *Flavobacteria* in Antarctic seawaters have been reported previously, and *Flavobacteria* are known for their strong associations with the algal bloom following the ice melt during summer time in Antarctica (Abell and Bowman 2005, Kirchman 2002). In this study, the concentration of Chl a at the surface ranged from 0.6 to 4.5 (mg/m^3), representing numerous stages of the algal bloom in the Ross Sea. In surface water, the increase in diverse complex organic matter derived from primary production by algae may provide increased growth opportunities to *Flavobacteria*, which express transporters that target biopolymers (e.g., polysaccharides, proteins, and proteoglycan) and degradative enzymes (e.g., glycosyl hydrolases and peptidases), thereby facilitating the utilization of complex substrates (Williams et al. 2013). In addition, members of *Flavobacteria* are known to encode proteorhodopsins and proteins involved in protection against damage caused by oxidative stress resulting from high solar irradiance in aerobic surface waters (Williams et al. 2013). These physiological traits of *Flavobacteria* may explain the high abundance of this group in surface water and enable them to play important roles in biogeochemical cycling in marine environments, functioning as heterotrophic degraders of high-molecular-weight compounds, particularly in surface waters (Abell and Bowman 2005, Claire Horner-Devine et al. 2003, Kirchman 2002, Williams et al. 2013).

In contrast, the relative abundance of *Gammaproteobacteria* increased as the depth increased, showing a negative correlation with *Bacteroidetes*. Regardless of the total proportion, the relative abundances of some phyla, such as *Deltaproteobacteria*, *Betaproteobacteria*, SAR406, *Actinobacteria*, *Verrucomicrobia*, *Chloroflexi*, *Gemmatimonadetes*, and *Planctomycetes*, which were rarely recovered in surface waters, were higher in deep waters. The high proportions of *Deltaproteobacteria*, *Actinobacteria*, *Verrucomicrobia*, *Chloroflexi*, and *Planctomycetes* along the depth were consistent with previous results described by Sunagawa et al. (2015), and SAR406 has been frequently recovered in deep waters (Fuhrman and Davis 1997, López-García et al. 2001, Signori et

al. 2014). Thus, these bacterial groups are likely to contribute to chemosynthetic production via versatile metabolic pathways using inorganic compounds. These bacterial lineages have shown to be frequently recovered in surface sediments (Kouridaki et al. 2010, Parkes et al. 2014, Polymenakou et al. 2009, Zinger et al. 2011), implying that the physical interconnection between deep waters and surface sediments results in similar environments that act as a selection pressure for bacterial groups.

2.4.3 Vertical and spatial distribution of bacterial composition in marine sediments of the Ross Sea, Antarctica

Highly diverse bacterial populations (40 bacterial phyla with more than 1% in relative abundance of total bacterial communities throughout the aerobic samples) compared to twelve phyla recovered from cloning of 16S rRNA genes in the Ross Sea was revealed (Baldi et al. 2010). Distinct bacterial communities between aerobic sediments and anaerobic sediments were revealed indicating the oxygen availability than geological location was the most important factor for shaping the bacterial communities. In the aerobic surface sediments, *Gammaproteobacteria*, *Deltaproteobacteria*, *Bacteroidetes*, and *Planctomycetes* dominated consistently with the previous results from other areas such as the Arctic, Antarctic, Eastern Mediterranean Sea, and Northeastern Pacific Ocean (Baldi et al. 2010, Bowman and McCuaig 2003, Kouridaki et al. 2010, Li et al. 2009, Park et al. 2011, Parkes et al. 2014, Polymenakou et al. 2009). These findings may indicate that members of the predominant taxa are well-adapted to the surficial layer of marine sediments regardless of geographical location.

A clear down-core shift in bacterial community composition from *Deltaproteobacteria*, *Gammaproteobacteria*, *Planctomycetes*, *Bacteroidetes*, *Acidobacteria*, and *Chlorobi* in the oxic surface sediment to Ca. *Atribacteria* JS1,

Chloroflexi, *Actinobacteria*, *Betaproteobacteria*, OP8 at and below 20 cmbsf were revealed. Analysis of bacterial communities until approximately 4 m in depth revealed that the community was largely clustered into three groups consistently with the change in environmental variables. The physiochemical properties across the sediments seemed to be closely related to centimeter-scale stratigraphic variability in the bacterial community. Moreover, the bacterial at 0 cmbsf was clearly distinct from that of anoxic sediments, potentially owing to differences in oxygen availability. Since the concentration of oxygen was not analyzed in this study and the depth of oxygen penetration into marine sediments differs significantly among regions, the effect of oxygen for shaping the microbial communities is not measurable (D/'Hondt et al. 2015, Revsbech et al. 1980, Walker 2005). Furthermore, the sediments beneath the Ross Ice Shelf were oxic with no dramatic shifts in redox states or carbon availability through the entire sediment to 1.5 mbsf (Carr et al. 2013). However, considering that dissolved oxygen penetrates the entire sediment column where both the mean sediment accumulation rate and total sediment thickness are low, the redox state of sediments in this study seemed to be different with that from sediments beneath the Ross Ice Shelf which persisted throughout the Holocene and thus, the sediments deposition was different with that from open area of the Ross Sea (Carr et al. 2013, D/'Hondt et al. 2015). Instead, since the study area is biologically productive and sedimentation rate is known to be high, it was presumable that oxygen in the sediment of this study is not expected to be present below a few centimeters of sediments likely with that from other oceanic sediments in regions (Arrigo and van Dijken 2004, Cai and Sayles 1996, D/'Hondt et al. 2015, Walker 2005). Bacterial community distinction at 120 cmbsf was well matched with the dramatic changes in environmental factors, such as total organic carbon, and water contents and the source of sediments.

OTU of *Atribacteria* JS1 lineage dominated throughout the anoxic sediments. The JS1 lineage is known to be one of the most predominant bacterial groups in various subseafloor sediments, especially strictly anoxic organic-rich environments or gas hydrate-containing sediments. To understand the metabolic potential of this predominant but

uncultivated group, culture-independent approaches such as isotope enrichment studies and genomics have been applied. Stable isotope probing and enrichments studies of marine sediments revealed the incorporation of acetate and glucose into members of JS1 suggesting the heterotrophic metabolism (Dodsworth et al. 2013, Webster et al. 2006b). Single-cell amplified genome (SAG) analysis or metagenome approaches revealed the lack of sugar fermentation pathways and capacity to catabolize the organic acids such as propionate and acetate via the methylmalonyl-CoA pathway (Nobu et al. 2015). For better understanding of the metabolic potential and ecological functions of this group in the Ross Sea, further studies seemed to be needed.

CHAPTER 3. Diversity and physiological characteristics of culturable bacteria from marine sediments of Ross Sea, Antarctica

3.1 Introduction

The fraction of benthic bacteria may account for 10% to 30% of the Earth's total biomass and approximately 70% of the global prokaryotic biomass (Li et al. 2009, Whitman et al. 1998). In addition, benthic bacterial communities in the ocean play significant roles in biogeochemical cycles and in the remineralization of organic materials (Li et al. 2009, Ravenschlag et al. 2001).

Among the diverse mechanisms that enable microorganisms to survive and thrive in cold environments are adaptive alterations in cellular proteins and lipids (Gerday et al. 2000, Russell 1998). In polar environments, extracellular enzymes secreted by cold-adapted microorganisms play a crucial ecological role in nutrient cycling (Staley and Herwig 1993, Vazquez et al. 2004). These cold-active enzymes are also highly valuable because of their potential biotechnological applications that reduce the energy required for chemical processes (Groudieva et al. 2004) and the potential use of exopolysaccharides from polar fungi and bacteria as cryoprotectants (Kim and Yim 2007, Selbmann et al. 2002). Thus, the ability of the newly isolated psychrophilic or psychrotolerant strains to produce a broad spectrum of cold-active enzymes and other materials is of great interest for both fundamental research and a broad range of industrial, agricultural, and medical applications (Yu et al. 2011).

New molecular techniques have enabled cultivation-independent approaches to be widely adopted in research into microbial diversity (Bai et al. 2006). Nonetheless, previous studies on benthic bacterial communities by culture-independent methods revealed that many of the OTUs detected had low similarities with known strains, indicating their unknown functions and physiologies (Bowman and McCuaig 2003, Polymenakou et al. 2005, Webster et al. 2006b). This observation highlighted the necessity for cultivation, to isolate previously undescribed strains and to understand their physiological characteristics. Determination of the physiological characteristics of

culturable microorganisms can provide information on their ecological roles of at least some members of the microbial community (Jiang et al. 2006). In addition, the culture-based approach is appropriate for assembling a collection of microorganisms for biochemical, genetic, and physiological studies (Jiang et al. 2006).

In this study, the taxonomic affiliations of bacterial isolates recovered from the sediments of Ross Sea, Antarctica were presented to gain insights into the nature of the culturable bacterial diversity. Additionally, the physiological characteristics of the obtained bacteria were investigated.

3.2 Materials and Methods

Samples and isolation of bacterial strains

Marine sediments were collected in 2011 from three sites in Ross Sea, Antarctica by a box corer (LYM110206-01), dredge (LYM110208-01), or grab sampler (LYM110211-02) (Table 2-2). The samples were suspended in 20% glycerol and preserved at -80°C until use. For cultivation of bacterial isolates, the samples were serially diluted up to 10^{-3} in sterilized sea water and 100-µl aliquots of the diluted sample suspensions were spread on four kinds of medium followed by incubation at 10°C for 12–22 days. The media used in these cultivations were: (1) marine agar (MA, Difco, USA), (2) 1/10 diluted marine agar (0.1× MA), (3) marine R2A (MR2A, Difco), and (4) 1/10 diluted marine R2A (0.1× MR2A) After incubation, colonies from the agar plates were picked on the basis of their morphology and subcultured in fresh agar medium until pure isolates were obtained. Pure cultures of the bacterial isolates were deposited in the Polar and Alpine Microbial Collection (PAMC) (Lee et al. 2012).

Examination of physiological characteristics

Cell suspensions cultured in marine broth medium (Difco) at 10°C and with shaking at 120 rpm were used to investigate the physiological characteristics of the respective isolates. Growth temperature and the production of extracellular protease, lipase, and exopolysaccharides were determined using the replica plating methods of Lee *et al.*.(Lee et al. 2012). The temperature-dependent growth response was studied by replica plating of cell suspensions with a 96-pin replicator (VP-408B, V&P Scientific, USA) followed by a 7-day incubation at 4°C, 10°C, 15°C, 20°C, 25°C, 30°C, or 37°C. Growth was evaluated by scoring the size and turbidity of the colony, as described by Lee *et al.*. (Lee et al. 2012). Protease and lipase secretion was examined by replica plating of cell suspensions onto 0.1× MA plates supplemented with 1% skim milk (Difco) or 1% tributyrate (Sigma, USA), respectively. The plates were incubated for 7 days at 10°C and 20°C, respectively. Enzyme secretion was scored based on the ratio of colony size to the width of the clear zone surrounding the colony. Exopolysaccharide (EPS) production was recognized by the formation of ropy colonies (Macura and Townsley 1984) at 20°C. The color of the isolates was determined visually.

Identification of bacterial isolates

Bacterial strains were identified based on the sequence similarity and phylogenetic analysis of 16S rRNA gene sequences. Bacterial genomic DNA was extracted using the LaboPass tissue mini kit (Cosmogenetech, Korea). The 16S rRNA gene was PCR-amplified with two universal primers, 27F and 1492R as described by Lane (Lane 1991). PCR was carried out using the method described by Lee et al. (Lee et al. 2012). PCR products were purified using the LaboPass PCR purification kit (Cosmogenetech) and sequenced with the same primers used for amplification. The sequence of the 16S rRNA gene was compared with those of type strains available in the EzTaxon-e database

(Kim et al. 2012a) to find closely related species and to choose reference sequences for the phylogenetic analyses. Phylogenetic trees were reconstructed by the neighbor-joining method (Saitou and Nei 1987) based on the distance matrix generated according to Kimura's two-parameter model (Kimura 1980) and using phydit ver. 3.2 (http://plaza.snu.ac.kr/~jchun/phydit/). The confidence level of the tree topology was evaluated by bootstrap analysis using 1,000 sequence replications. The species affiliation of a bacterial isolate was determined when the isolate formed a monophyletic group with the reference species and based on a 98.65% or higher similarity (Kim et al. 2014).

3.3 Results

Phylogenetic identification of the isolates

Sixty-three isolates, affiliated with *Actinobacteria*, *Bacteroidetes*, *Alphaproteobacteria*, and *Gammaproteobacteria*, were obtained from three sediment samples (Table 3-1). Two of the isolates belonged to the phylum *Actinobacteria*, represented by the genera *Cryobacterium*, and *Marisediminicola*; 36 isolates belonged to the phylum *Bacteroidetes*, represented by the genera *Algibacter* (1 isolate), *Flavobacterium* (15 isolates), *Lacinutrix* (1 isolate), *Polaribacter* (3 isolates), *Psychroserpens* (8 isolates), *Winogradskyella* (7 isolates), and the unidentified genus of family *Flavobacteriaceae* (1 isolate); 17 isolates of the class *Alphaproteobacteria* belonged to the genera *Loktanella* (10 isolates), *Octadecabacter* (2 isolates), *Roseovarius* (1 isolate), and *Sulfitobacter* (4 isolates); and 7 to the genus *Psychrobacter* and 1 member of the genus *Shewanella* of *Gammaproteobacteria*. Among these isolates, the largest groups in terms of the number of isolates recovered were those belonging to the genera *Flavobacterium* (15 isolates), *Loktanella* (10 isolates), *Psychroserpens* (8 isolates), and *Winogradskyella* (7 isolates) and *Psychrobacter* (7 isolates). Isolates affiliated with the genera *Psychroserpens* and *Sulfitobacter* were obtained from all three sediments. The

overall similarity of the isolates to known type strains ranged from 95.7% to 100% and 31 isolates (49.2 % over total strains) had ≤98.65% 16S rRNA gene sequence similarity with the nearest type strains (Table 3-1).

Physiological characteristics

The growth temperature of the 63 isolates and their production of extracellular enzymes, such as protease and lipase, and polymers were determined. Most of the isolates (86%) produced apricot, beige, yellow, or red pigments readily recognizable with the unaided eye (Table 3-1). As the temperature increased from 4°C to 10°C, the number of growing strains increased from 39 to 63 and then gradually decreased from 15°C. Most of the isolates grew well between 10°C and 15°C but only three strains could grow at 37°C (Fig. 3-1).

Extracellular protease activities were detected in 29 isolates, belonging to the genera *Cryobacterium*, *Flavobacterium*, *Polaribacter*, *Psychroserpens*, *Winogradskyella*, *Loktanella*, *Sulfitobacter*, and *Psychrobacter* (Table 3-1). Strains belonging to the genera *Flavobacterium* (45%), *Psychrobacter* (19%), and *Loktanella* (13%) accounted for a large proportion of the extracellular protease producers (Fig. 3-2a). Proteolysis at 10°C and 20°C was detected in 27 and 21 isolates, respectively (Table 3-1). Nineteen isolates showed extracellular protease activities both at 10°C and 20°C and higher extracellular protease activity was observed at 10°C than 20°C. In particular, 21 isolates, belonging to the phylotypes *Cryobacterium psychrotolerans*, *Flavobacterium* sp., *Psychroserpens* sp., *Winogradskyella* sp., *Loktanella salsilacus*, *Loktanella* sp., and *Psychrobacter luti*, (Table 3-1 and Fig. 3-4), had high extracellular protease activity (score 3≥) at 10°C (Fig. 3-3a).

Sixteen isolates, belonging to the genera *Flavobacterium*, *Polaribacter*, *Psychroserpens*, *Winogradskyella*, *Loktanella*, *Sulfitobacter*, *Psychrobacter*, an

Table 3-1. Taxonomic assignments and physiology of the bacterial isolates

Species name	Similarity (%)	Temperature range (°C)	Protease*		Lipase*		EPS production	Pigment characteristics	Sample	PAMC No.	Accession No.
			10°C	20°C	10°C	20°C					
Actinobacteria											
Cryobacterium	99.3	10-15	3	2	0	0	No	yellowish	Sed3	27129	KJ475137
Marisediminicola	100.0	4-30	0	0	0	0	No	apricot	Sed3	27228	KJ475136
Bacteroidetes											
Algibacter sp.	97.2	4-15	0	0	0	0	No	yellow	Sed1	27237	KJ475138
Flavobacteriaceae	95.9	4-30	0	0	2	2	No	yellow	Sed1	27105	KJ475139
Flavobacterium	100.0	10-37	2	2	2	3	No	yellow	Sed3	27103	KJ475144
Flavobacterium	99.6	4-25	0	0	0	0	Yes	yellow	Sed3	27227	KJ475154
Flavobacterium sp.	99.4	10-30	4	1	0	1	No	yellow	Sed3	27098	KJ475140
Flavobacterium sp.	98.9	4-15	0	2	0	0	No	reddish	Sed3	27099	KJ475141
Flavobacterium sp.	99.2	10-15	3	0	0	0	Yes	yellow	Sed3	27101	KJ475142
Flavobacterium sp.	99.0	10-20	3	1	0	0	Yes	yellow	Sed3	27102	KJ475143
Flavobacterium sp.	99.1	4-30	3	3	2	2	No	yellow	Sed3	27115	KJ475145
Flavobacterium sp.	99.1	4-30	4	3	0	0	No	yellow	Sed3	27122	KJ475146
Flavobacterium sp.	99.1	4-20	3	0	0	0	No	yellow	Sed3	27123	KJ475147
Flavobacterium sp.	99.1	4-20	3	0	0	0	Yes	yellow	Sed3	27124	KJ475148
Flavobacterium sp.	99.1	10-20	4	2	0	0	No	yellow	Sed3	27125	KJ475149
Flavobacterium sp.	99.2	4-20	3	0	0	0	No	yellow	Sed3	27131	KJ475150

Table 3-1. To be continued

Species name	Similarit (%)	Temperature range (°C)	Protease* 10°C	Protease* 20°C	Lipase* 10°C	Lipase* 20°C	EPS production	Pigment characteristics	Sample ID§	PAMC No.	Accession No.
Flavobacterium sp.	99.1	4-20	4	3	0	0	Yes	yellow	Sed3	27133	KJ475151
Flavobacterium sp.	99.2	4-20	4	3	0	0	No	yellow	Sed3	27134	KJ475152
Flavobacterium sp.	99.1	4-20	4	2	0	0	Yes	yellow	Sed3	27207	KJ475153
Lacinutrix sp.	97.5	4-10	0	0	0	0	No	yellow	Sed1	27137	KF977035
Polaribacter sp.	97.9	10-30	0	0	0	0	Yes	yellow	Sed1	27095	KJ475155
Polaribacter sp.	97.9	4-30	0	0	2	2	Yes	yellow	Sed1	27096	KJ475156
Polaribacter sp.	98.1	10-15	1	0	0	0	Yes	reddish	Sed3	27100	KJ475157
Psychroserpens	99.2	4-30	0	0	0	0	Yes	yellow	Sed2	27240	KJ475165
Psychroserpens sp.	97.6	10-20	0	0	0	0	No	reddish	Sed1	27104	KJ475158
Psychroserpens sp.	97.8	4-30	0	0	1	1	Yes	reddish	Sed1	27106	KJ475159
Psychroserpens sp.	97.0	10	0	0	0	0	No	greenish	Sed3	27130	KJ475160
Psychroserpens sp.	97.3	4-20	4	2	0	0	Yes	yellow	Sed1	27206	KJ475161
Psychroserpens sp.	97.7	10-15	3	3	0	0	No	reddish	Sed1	27216	KJ475162
Psychroserpens sp.	97.1	10-15	0	0	0	0	No	yellow	Sed1	27220	KJ475163
Psychroserpens sp.	97.8	4-30	0	0	0	0	No	reddish	Sed1	27238	KJ475164
Winogradskyella sp.	98.5	4-30	0	0	1	2	Yes	apricot	Sed1	27097	KJ475166
Winogradskyella sp.	98.5	4-30	0	0	2	1	Yes	apricot	Sed1	27107	KJ475167
Winogradskyella sp.	98.5	10-30	0	0	0	0	Yes	reddish	Sed1	27217	KJ475171
Winogradskyella sp.	98.6	10-15	0	0	0	0	No	reddish	Sed1	27221	KJ475172

Table 3-1. To be continued

Species name	Similarit (%)	Temperature range (°C)	Protease* 10°C	Protease* 20°C	Lipase* 10°C	Lipase* 20°C	EPS production	Pigment characteristics	Sample ID§	PAMC No.	Accession No.
Winogradskyella sp.	97.0	10-20	3	1	0	0	No	yellow	Sed3	27136	KJ475168
Winogradskyella sp.	97.1	10-15	0	0	0	0	No	yellow	Sed3	27139	KJ475169
Winogradskyella sp.	97.0	10-37	0	0	0	0	Yes	yellow	Sed3	27140	KJ475170
Alphaproteobacteria											
Loktanella	99.9	4-30	3	0	0	0	Yes	beige	Sed3	27121	KJ475175
Loktanella	100.0	4-30	0	0	0	0	No	beige	Sed3	27229	KJ475181
Loktanella sp.	95.7	4-30	0	0	1	0	Yes	beige	Sed1	27117	KJ475173
Loktanella sp.	95.7	4-30	0	0	1	0	Yes	beige	Sed1	27118	KJ475174
Loktanella sp.	97.0	10	0	0	0	0	No	beige	Sed1	27126	KJ475176
Loktanella sp.	97.0	4-20	3	3	0	0	No	beige	Sed3	27132	KJ475177
Loktanella sp.	96.7	4-20	4	2	0	0	No	cream	Sed3	27135	KJ475178
Loktanella sp.	96.8	10-15	0	0	0	0	No	beige	Sed1	27138	KJ475179
Loktanella sp.	96.4	10-15	3	2	0	0	No	beige	Sed1	27223	KJ475180
Loktanella sp.	97.1	10-25	0	0	0	0	No	beige	Sed3	27241	KJ475182
Octadecabacter	99.9	10-25	0	0	0	0	No	beige	Sed1	27224	KJ475183
Octadecabacter	99.9	10-25	0	0	0	0	No	beige	Sed1	27225	KJ475184
Roseovarius sp.	97.0	10-20	0	0	0	0	No	beige	Sed3	27236	KJ475185
Sulfitobacter	100.0	4-30	2	0	0	0	No	beige	Sed2	27120	KJ475187
Sulfitobacter	99.6	4-30	2	2	0	2	No	cream	Sed2	27109	KJ475186

Table 3-1. To be continued

Species name	Similarit (%)	Temperature range (°C)	Protease* 10°C	Protease* 20°C	Lipase* 10°C	Lipase* 20°C	EPS production	Pigment characteristics	Sample ID§ No.	PAMC No.	Accession
Sulfitobacter sp.	98.1	10	0	0	0	0	No	beige	Sed1	27222	KJ475188
Sulfitobacter sp.	98.1	10-15	0	0	0	0	No	beige	Sed1	27232	KJ475189
Gammaproteobacteri											
Psychrobacter luti	99.8	4-30	4	3	2	1	Yes	cream	Sed3	27116	KJ475194
Psychrobacter luti	100.0	4-30	2	2	0	0	No	beige	Sed2	27119	KJ475195
Psychrobacter	99.9	4-30	0	1	2	0	No	cream	Sed2	27108	KJ475190
Psychrobacter	99.9	4-30	0	0	1	2	No	cream	Sed2	27110	KJ475191
Psychrobacter	99.8	4-30	0	0	1	2	No	cream	Sed2	27111	KJ475192
Psychrobacter	99.9	4-30	1	0	1	1	No	cream	Sed2	27112	KJ475193
Psychrobacter	99.9	10-30	0	0	0	0	No	cream	Sed2	27239	KJ475196

* The scores (from 1 to 4) represent the degree of production, with a higher number implying better production of protease and lipase.

§ Sed1 is LYM110206-01, Sed2 is LYM110208-01, and Sed3 is LYM110211-02 in Table 2-16.

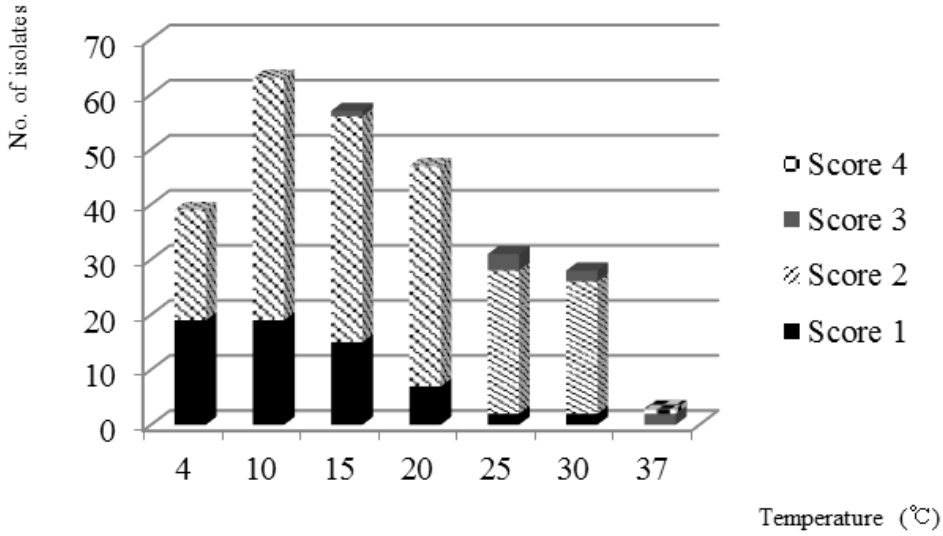


Figure 3-1. Effect of temperature on bacterial growth.

The scores (from 1 to 4) represent the degree of growth, with a higher number implying better growth.

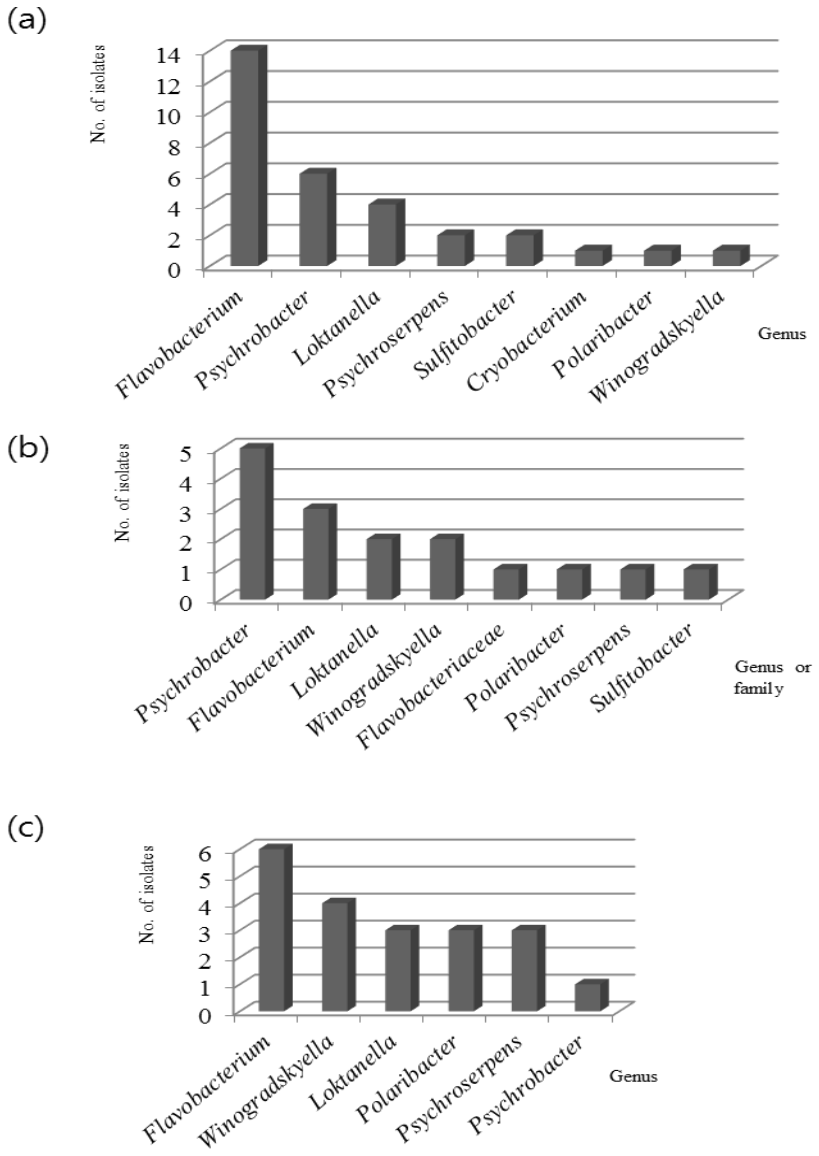


Figure 3-2. Abundance of isolates, according to genus or family, that produce extracellular protease (a), lipase (b), and exopolysaccharides (c).

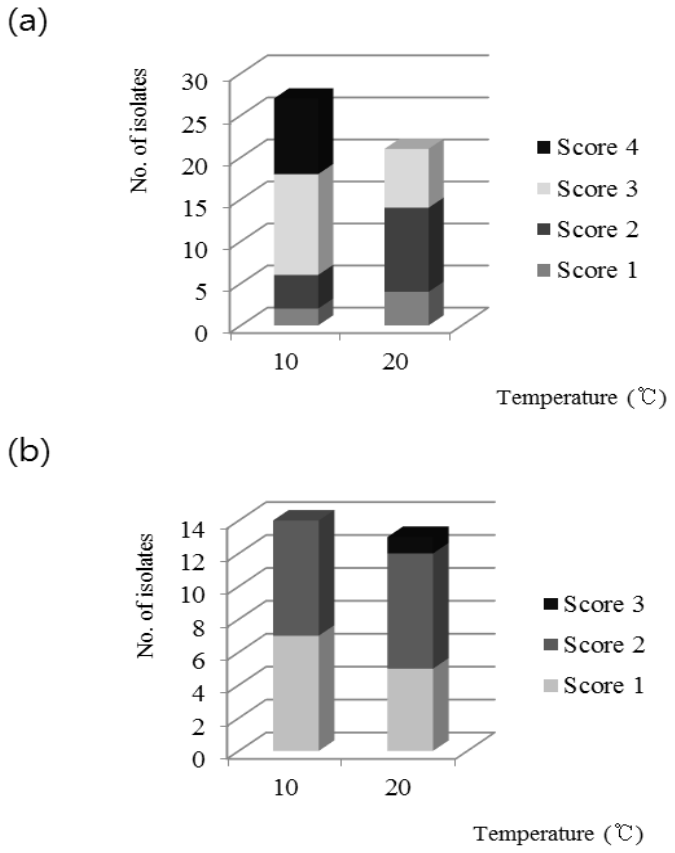


Figure 3-3. Frequency of bacterial isolates with extracellular protease (a) and lipase (b) activities. The scores (from 1 to 4) represent the degree of production, with a higher number implying better production.

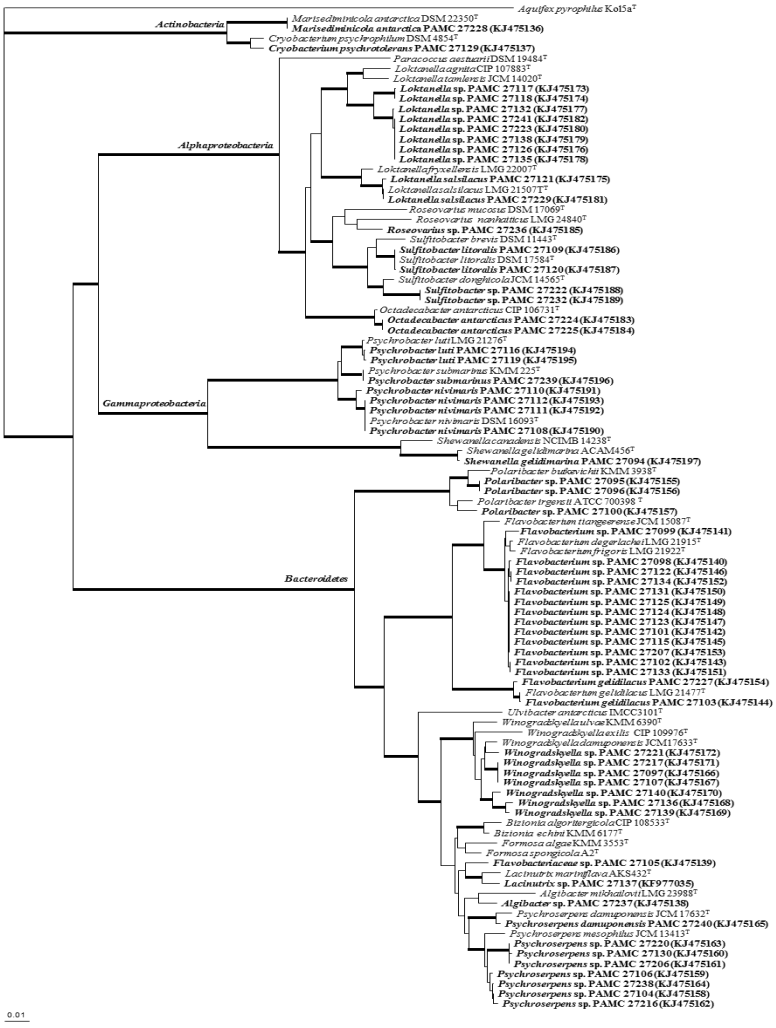


Figure 3-4. Neighbor-joining tree of isolates with closely related reference species, based on the 16S rRNA gene sequences. Representative isolates for each phylotype are indicated by bold letters, and branches supported by high bootstrap values (>70%) as thick lines. Bar, 1 nucleotide substitutions per 100 nucleotide

unidentified genus of the family *Flavobacteriaceae*, had extracellular lipase activity, with strains of the genera *Psychrobacter* (31%), *Flavobacterium* (19%), *Loktanella* (13%), and *Winogradskyella* (13%) accounting for a large proportion of the extracellular lipase producers (Fig. 3-2b). Fourteen isolates exhibited enzyme activity at 10°C and 13 isolates exhibited enzyme activity at 20°C (Table 3-1). Among the 11 isolates with extracellular lipase activity at both 10°C and 20°C, score-based enzyme activity was higher at the higher temperature (Fig. 3-3b). The maximum score of lipase secretion was 3 at 20°C by *Flavobacterium gelidilacus* PAMC 27103 (Table 3-1 and Fig. 3-4).

The 20 isolates that produced exopolysaccharides at 20°C were affiliated with the genera *Flavobacterium*, *Polaribacter*, *Psychroserpens*, *Winogradskyella*, *Loktanella*, and *Psychrobacter* (Table 3-1). Among them, strains belonging to the genera *Flavobacterium* (30%) were the most abundant followed by *Winogradskyella* (20%), *Loktanella* (15%), *Polaribacter* (15%), and *Psychroserpens* (15%) (Fig. 3-2c). Approximately 43 isolates (70%) produced at least one extracellular enzyme or exopolysaccharide, and the isolate, PAMC 27116, with 99.2% similarity to *Psychrobacter luti*, produced three of them (Table 3-1).

3.4 Discussion

In the sediments from Ross Sea, Antarctica, both the spatial distribution of the organic matter composition and bacterial densities have been reported (Baldi et al. 2010, Fabiano and Danovaro 1998, Fabiano and Pusceddu 1998, Pusceddu et al. 2000). The bacterial community of Ross Sea sediments, as determined by T-RFLP analysis, revealed the predominance of bacteria belonging to *Gammaproteobacteria*, *Deltaproteobacteria*, *Bacteroidetes*, and *Acidobacteria* (Baldi et al. 2010). To the best of our knowledge, this is the first report on the diversity and physiological characteristics of culturable bacteria in Ross Sea sediments. The culturable bacteria present in the samples were affiliated with 16

genera and many of them (70%) produced extracellular protease, lipase, and/or exopolysaccharides. Most of the organic matter in benthic communities is produced as high-molecular-weight polymeric compounds, which before they can be incorporated into microbial cells must be degraded by a series of extracellular hydrolytic enzymes (Yu et al. 2011). The finding that a large proportion of the isolates produced at least one type of extracellular enzymes points to the significant roles by benthic microorganisms in the biogeochemical cycles of Ross Sea sediments. In addition, analyses of temperature-dependent growth showed that the number of isolates capable of growing at 10°C was highest, indicating an adaption at the stable, low temperature of the bottom sediments in the SO (Helmke and Weyland 2004).

The predominance of the phylum *Bacteroidetes* is known in the previous studies of bacterial communities from marine sediments of Antarctica (Baldi et al. 2010, Bowman and McCuaig 2003). In this study, 59% of total recovered isolates belonged to the phylum *Bacteroidetes*. Among them, the large percentage (67%, 29 isolates out of 43) of extracellular-material-producing strains were affiliated with the genera *Flavobacterium*, *Polaribacter*, *Psychroserpens*, *Winogradskyella*, and unidentified genus of the family *Flavobacteriaceae* and it is consistent with the well-established finding of biopolymer hydrolysis by this phylum (Kirchman 2002). In addition, extracellular protease or lipase production by isolates assigned to the genera *Cryobacterium*, *Loktanella*, *Psychrobacter*, and *Sulfitobacter*, belonging to *Actinobacteria*, *Alphaproteobacteria*, and *Gammaproteobacteria*, reflects the importance of these strains in the hydrolysis of organic constituents.

Twenty-one isolates, affiliated with the genera *Flavobacterium*, *Polaribacter*, *Psychroserpens*, *Winogradskyella*, *Loktanella*, *Sulfitobacter*, and *Psychrobacter* produced at least two extracellular proteases, lipases, and/or exopolysaccharides. Of these isolates, members of the genus *Flavobacterium* (33%) clearly dominated. The strains of the genus *Flavobacterium* are known for their specialized roles in the uptake and degradation of the

high-molecular-mass fraction of dissolved organic matter and in remineralization processes, both in freshwater and in marine ecosystems (McCammon and Bowman 2000). They have been frequently reported in oligotrophic and eutrophic Antarctic freshwater, terrestrial samples, and marine sediments, suggesting their wide diffusion in Antarctica and their ecological roles in macromolecule hydrolysis (Humphry et al. 2001, McCammon et al. 1998, McCammon and Bowman 2000) (Michaud et al. 2012, Van Trappen et al. 2003, Van Trappen et al. 2004, Van Trappen et al. 2005, Yi et al. 2005, Yi and Chun 2006).

Although culturable bacteria do not fully represent the bacterial communities of the environment, many candidate novel species with <98.65% similarity to known strains, were recovered in this study. This result demonstrates that relatively simple cultivation methods can be used to isolate as-yet-undescribed taxa and they can be used for expanding the knowledge on the unknown functions and physiologies of bacterial OTUs obtained by molecular techniques. In addition, our finding that culturable bacteria produce cold-active enzymes indicates that the isolated strains contribute to the hydrolysis of major organic constituents and are therefore involved in carbon and nitrogen cycling at the low temperature of sediments of Ross Sea, Antarctica.

CHAPTER 4. Genomic insight to predominant *Candidatus Atribacteria* JS1 lineage in marine sediments of the Ross Sea, Antarctica

4.1 Introduction

Many previous studies on the bacterial diversity of subseafloor revealed the majority of microorganisms belong to previously unidentified and uncultured groups, called candidate divisions or phyla, making it difficult to infer their metabolic activities and ecological functions (Corinaldesi 2015b, Eloe-Fadrosh et al. 2016, Handelsman 2004, Inagaki et al. 2006, Macaulay and Voet 2014, Nobu et al. 2015, Orcutt et al. 2011, Rinke et al. 2013). Thus, understanding the metabolism and functions of candidate divisions in the subseafloor biosphere is a major challenge and will provide substantial insights into the roles of microbes in global biogeochemical cycles.

Candidatus (Ca.) *Atribacteria* is a candidate phylum that was proposed to include members of the OP9 and JS1 lineages (Nobu et al. 2015) The JS1 lineage is one of the most predominant bacterial groups in various subseafloor sediments, especially strictly anoxic organic-rich environments or gas hydrate-containing sediments, comprising over 50% of bacterial communities in some marine sediments (Carr et al. 2015, Inagaki et al. 2006, Nobu et al. 2015, Nunoura et al. 2016, Orcutt et al. 2011, Parkes et al. 2014, Webster et al. 2006a). To elucidate the physiology of the JS1 lineage, culture-independent approaches such as stable isotope enrichment studies and genomics have been applied. Stable isotope probing and enrichments studies suggested the heterotrophic metabolism of JS1 (Webster et al. 2006b, Webster et al. 2011). Single-cell amplified genome (SAG) analysis or metagenome approaches from marine sediments from Aarhus Bay, Etoliko Lagoon, and Adélie Basin, biofilms from a terephtalate-degrading reactor, and monimolimnion of meromictic Sakinaw Lake also indicated that members of JS1 are heterotrophic anaerobes that lack respiratory capacity (Carr et al. 2015, Dodsworth et al. 2013, Lloyd et al. 2013, Nobu et al. 2015, Rinke et al. 2013). The high coverage JS1 genomes recovered from meromictic Sakinaw Lake and mesophilic bioreactor revealed the lack of sugar fermentation pathways and capacity to catabolize the organic acids such as propionate and acetate via the methylmalonyl-CoA pathway (Nobu et al. 2015).

However, the low coverage JS1 genomes from the marine sediments (<22%) did not allow the detailed description of their metabolic potential in spite of the predominance in those environments and it remained the metabolism and ecological functions of the JS1 lineage in marine sedimnets largely unknown (Carr et al. 2015, Lloyd et al. 2013, Nobu et al. 2015, Rinke et al. 2013).

The vertical profiles of bacterial communities along a 3.96-m-long sediment core in the Ross Sea, Antarctica using 454 pyrosequencing of the 16S rRNA gene revealed JS1 lineage was one of the predominant bacterial groups throughout the sediments below the surface (Fig. 2-16a). To better understand the metabolic features and ecological functions of the JS1 lineage in this environment, genomes obtained by single cell sorting and amplification were analyzed.

4.2 Materials and Methods

Single-cell sorting, genome amplification, sequencing, and phylogenetic analysis

Based on bacterial community results, a sample from 40 cmbsf of DG12-GC06 with a high portion of the Ca. *Atribacteria* JS1 lineage was selected for single-cell sorting (Table 2-2 and Fig. 2-16a). Samples preserved in 20% glycerol at -80°C were centrifuged for 30 s at 9,300 × *g*, and 1 mL of supernatant was mixed with 100 μL of 100× TE buffer (pH 8.0), packed in dry ice, and sent to Bigelow Lab (East Boothbay, USA). Physical isolation of single cells was performed by fluorescent-activated cell sorting in a 384-well plate. After single-cell sorting, lysis of single cells and amplification of the single-cell genome by multiple displacement amplification (MDA) were performed. MDA product subsamples were used as a template in PCR for amplification of bacterial 16S ribosomal RNA genes using the primer sets 27F and 1492R (Lane 1991). Eighteen SAGs belonging to the JS1 lineage of *Atribacteria* were sequenced using a MiSeq sequencer system (Illumina) at Chun Lab (Korea). For 16S rRNA gene phylogenies, sequences obtained in

this study were aligned with those of *Atribacteria* retrieved from the SAGs and from metagenomic data sets and one dominant OTU of JS1 lineage obtained from pyrosequencing results using the RDP-II aligner (Cole et al. 2013). A phylogenetic tree was constructed using the neighbor-joining algorithm (Saitou and Nei 1987) with MEGA 6 (Tamura et al. 2013). The robustness of the tree topologies was assessed by bootstrap analyses based on 1,000 replications of the sequences.

Genome assembly and annotation

The quality trimming of sequencing reads and initial assemblies of 18 SAG sequences were performed using CLC Genomics Workbench (version 8.0) with default parameters. Average nucleotide identity (ANI) values were calculated to identify whether the SAGs were identical to each other using the dnadiff script of the MUMmer tool (version 3.0) (Kurtz et al. 2004). Single-cell genomes above 99% ANI were selected for further analysis (Dodsworth et al. 2013, Marshall et al. 2012). To construct the composite SAG (cSAG), designated RS JS1-cSAG, reads from 18 SAGs were pooled and assembled using CLC Genomics Workbench. Chimeric reads were removed by the jackknifing procedure (Marshall et al. 2012).

Assembled contigs were annotated with Rapid Annotation using Subsystem Technology (Aziz et al. 2008). Additional enzyme commission (EC) numbers for coding sequences were obtained from the KEGG Automatic Annotation Server (KAAS) (Moriya et al. 2007).

Genome completeness estimation

Genome completeness was estimated using the Conserved Single Copy Gene (CSCG) set, as previously described (Rinke et al. 2013). To identify CSCGs in RS JS1-

cSAG, protein sequences for PFAM matches were analyzed (Finn et al. 2016). The search for protein families using HMMER3 was performed using an online tool (http://pfam.xfam.org/). Proteins with the resulting best hits above precalculated HMM cutoffs were selected as CSCGs, and the completeness was estimated as the ratio of CSCGs to total CSCGs. Since the CSCGs were defined to occur only once in at least 90% of all genomes (n = 1,516), the number of total CSCGs was normalized to 90% (Rinke et al. 2013).

4.3 Results

Genome characteristics of single amplified genomes and phylogenetic diversity of *Atribacteria*

Eighteen SAGs of this lineage from a sediment sample (40 cmbsf) were recovered from single cell sorting and genome amplification. The size of each SAG ranged from 300 to 893 kb (Table 4-1). The SAGs shared 16S rRNA gene sequence similarity higher than 99.3% and had more than 99.2% 16S rRNA gene sequence similarity with c4865, the most predominant OTU of *Atribacteria* JS1. When 16S rRNA gene sequences from this study were compared with those retrieved from *Atribacteria* JS1 genomes from other habitats, SCGC_AD-561_N23 from Adélie Basin offshore Antarctica showed the highest similarity (96.1–97.3%), followed by ASPA from meromictic Sakinaw lake (94.0%). The 16S rRNA gene sequence similarity shared with that from genomes from other habitats, such as marine sediments from Aarhus Bay and Etoliko Lagoon and biofilm from a terephthalate-degrading reactor, was less than 94.0%. Phylogenies inferred from 16S rRNA gene sequences revealed that the *Atribacteria* JS1 lineage obtained in this study formed a monophyletic clade distinct from that of other habitats (Fig. 4-1).

Table 4-1. Genome characteristics of single amplified genomes (SAGs) from *Atribacteria*

SAG* ID	Total reads	# reads assembled	Total # contigs	N50 contig size (bp)	Max contig size (bp)	Assembly size (bp)	G+C contents (%)
AG183A05	3,981,694	3,533,168	112	10,286	35,313	604,612	33.4
AG183B19	4,432,318	3,961,380	122	6,200	21,868	504,934	36.9
AG183D19	5,005,792	4,467,434	108	8,012	32,374	458,549	34.3
AG183G20	3,345,186	2,916,062	156	7,062	24,763	664,301	33.8
AG183N14	3,525,408	3,112,864	72	6,598	18,407	300,007	34.9
AG183P08	4,114,504	3,642,932	74	8,880	26,459	369,086	37.7
AG183P14	4,527,826	4,067,252	99	6,589	23,709	435,436	33.7
AG183P18	4,239,382	3,607,000	136	13,521	64,226	751,925	31.3
AG183C13	3,478,788	3,388,806	548	5,036	25,965	667,274	35.3
AG183J17	2,786,182	2,727,664	407	1,461	25,397	335,029	36.0
AG183P03	3,010,548	2,940,758	924	2,185	36,505	893,508	35.1
AG183P13	2,987,058	2,917,108	450	3,980	32,609	456,911	35.6
AG183D10	3,785,158	3,656,556	452	1,348	30,580	361,811	36.6
AG183O15	3,515,540	3,419,538	561	3,857	35,945	657,801	35.5
AG183A11	2,843,454	2,757,232	723	3,433	19,535	763,790	34.9
AG183N04	3,343,960	3,112,864	756	3,323	32,857	826,550	35.5
AG183M21	3,741,560	3,629,894	633	3,269	26,741	653,913	35.1
AG183F16	3,139,342	3,050,738	611	3,072	26,157	675,445	35.0

Table 4-2. Genomic characteristics of *Atribacteria* JS1 lineage

Genome ID	NCBI accession	Habitat	Location	RAST by Nobu et al.[1]	RAST in this study[2]	Genome size (Mb)	Genome coverage (%)	Gene count in functional categories
RS JS1-cSAG		Marine sediment	Ross Sea, Antarctica		6666666.164	2.3	90	831
SCGC AD-561-N23		Marine sediment	Adelie Basin,		1476961.3	0.2		32
SL SAG co-assembly	AWNT00000000	Meromictic lake	Sakinow lake, Canada	6666666.54176	6666666.138	2.09	81	805
SCGC AAA252-M02	AQYX00000000	Meromictic lake	Sakinow lake, Canada		6666666.19	2.2		816
SL MG bin		Meromictic lake	Sakinow lake, Canada	6666666.54179	6666666.193	0.34	31	155
SCGC AB-164-G04	AQRY00000000	Meromictic lake	Sakinow lake, Canada	6666666.542	6666666.193	0.33	23	135
SCGC AAA255-E04	ASLT00000000	Meromictic lake	Sakinow lake, Canada	6666666.1	6666666.193	0.15		49
SCGC AAA255-G05	ASPA00000000	Meromictic lake	Sakinow lake, Canada	6666666.1	6666666.138	1.65		608
SCGC AAA255-N14	ASPC00000000	Meromictic lake	Sakinow lake, Canada	6666666.1	6666666.193	1.42		482
SCGC AB-164-A22	AQSW00000000	Meromictic lake	Sakinow lake, Canada	6666666.1	6666666.193	0.94		389
Aarhus Bay SAG I22	CDPM01000000	Marine sediment	Aarhus Bay, Baltic	6666666.945	6666666.162	1.04	22	300
Aarhus Bay SAG B17	CDPL01000000	Marine sediment	Aarhus Bay, Baltic	6666666.542	6666666.193	1.13	7	221
JGI 0000014-F07	ASOZ00000000	Marine sediment	Etoliko Lagoon,	6666666.542	6666666.193	0.32	8	92
JGI 0000079-L04	ASOY00000000	Terephthalate		6666666.542	6666666.193	0.95	25	434
TA biofilm MG bin		Terephthalate		6666666.542	6666666.193	2.01	86	688
JGI 0000059-I14	ASLS00000000	Terephthalate		6666666.542	6666666.193	0.49	33	173

[1] All assemblies are publicly available at RAST (http://rast.nmpdr.org/) by logging in with username and password “guest”

[2] Re-annotated assemblies for the comparison of RS JS1-cSAG and other *Atribacteria* JS1 lineages in RAST.

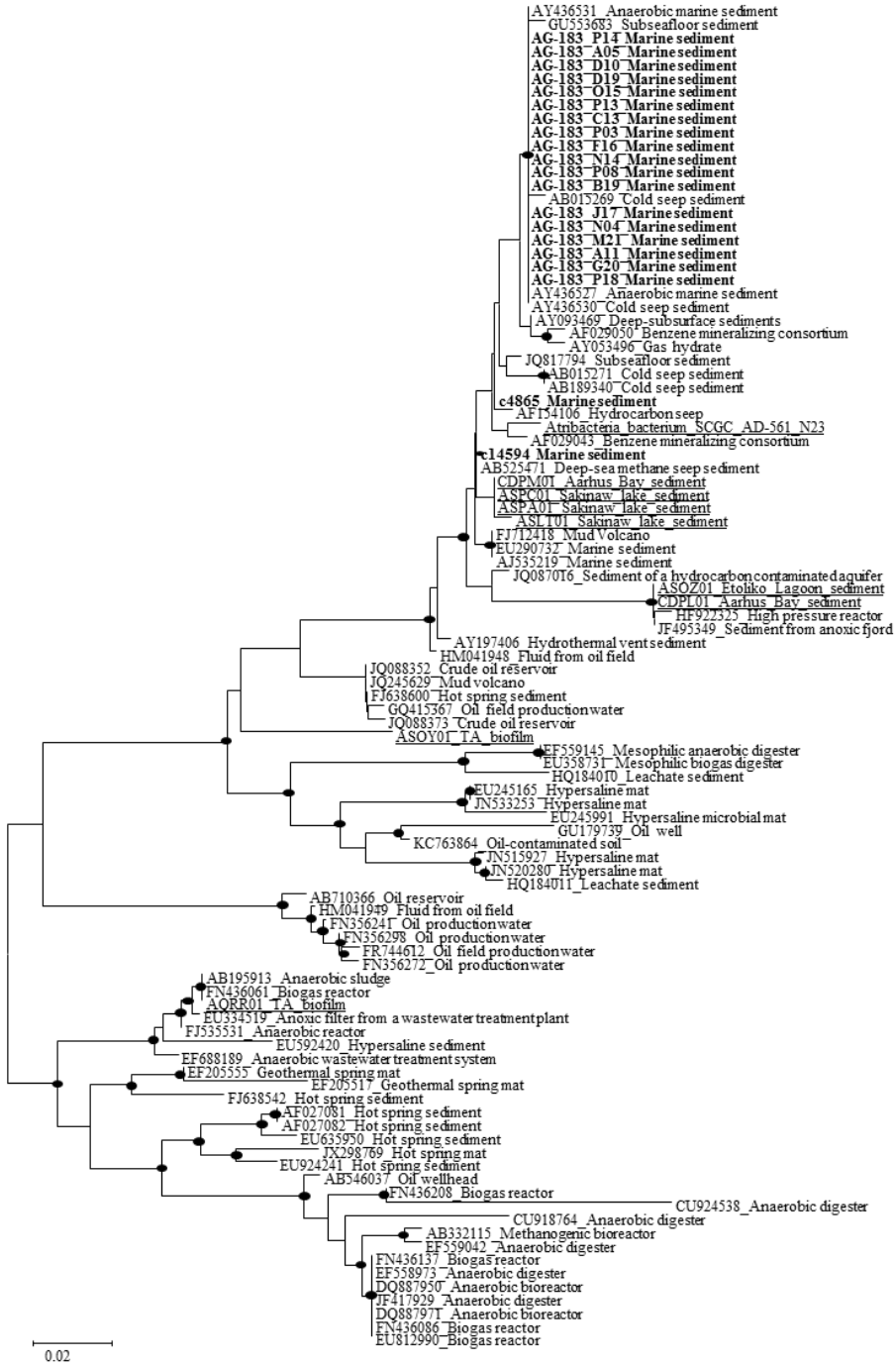

AY436531_Anaerobic marine sediment
GU553683_Subseafloor sediment
AG-183_P14_Marine sediment
AG-183_A05_Marine sediment
AG-183_D10_Marine sediment
AG-183_D19_Marine sediment
AG-183_O15_Marine sediment
AG-183_P13_Marine sediment
AG-183_C13_Marine sediment
AG-183_P03_Marine sediment
AG-183_F16_Marine sediment
AG-183_N14_Marine sediment
AG-183_P08_Marine sediment
AG-183_B19_Marine sediment
AB015269_Cold seep sediment
AG-183_J17_Marine sediment
AG-183_N04_Marine sediment
AG-183_M21_Marine sediment
AG-183_A11_Marine sediment
AG-183_G20_Marine sediment
AG-183_P18_Marine sediment
AY436527_Anaerobic marine sediment
AY436530_Cold seep sediment
AY093469_Deep-subsurface sediments
AF029050_Benzene mineralizing consortium
AY053496_Gas hydrate
JQ817794_Subseafloor sediment
AB015271_Cold seep sediment
AB189340_Cold seep sediment
c4865_Marine sediment
AF154106_Hydrocarbon seep
Atribacteria_bacterium_SCGC_AD-561_N23
AF029043_Benzene mineralizing consortium
c14594_Marine sediment
AB525471_Deep-sea methane seep sediment
CDPM01_Aarhus_Bay_sediment
ASPC01_Sakinaw_lake_sediment
ASPA01_Sakinaw_lake_sediment
ASLT01_Sakinaw_lake_sediment
FJ712418_Mud Volcano
EU290732_Marine sediment
AJ535219_Marine sediment
JQ087016_Sediment of a hydrocarbon contaminated aquifer
ASOZ01_Etoliko_Lagoon_sediment
CDPL01_Aarhus_Bay_sediment
HF922525_High pressure reactor
JF495349_Sediment from anoxic fjord
AY197406_Hydrothermal vent sediment
HM041948_Fluid from oil field
JQ088352_Crude oil reservoir
JQ245629_Mud volcano
FJ638600_Hot spring sediment
GQ415367_Oil field production water
JQ088373_Crude oil reservoir
ASOY01_TA_biofilm
EF559145_Mesophilic anaerobic digester
EU358731_Mesophilic biogas digester
HQ184010_Leachate sediment
EU245165_Hypersaline mat
JN533253_Hypersaline mat
EU245991_Hypersaline microbial mat
GU179739_Oil well
KC763864_Oil-contaminated soil
JN515927_Hypersaline mat
JN520280_Hypersaline mat
HQ184011_Leachate sediment
AB710366_Oil reservoir
HM041949_Fluid from oil field
FN356241_Oil production water
FN356298_Oil production water
FR744612_Oil field production water
FN356272_Oil production water
AB195913_Anaerobic sludge
FN436061_Biogas reactor
AQRR01_TA_biofilm
EU334519_Anoxic filter from a wastewater treatment plant
FJ535531_Anaerobic reactor
EU592420_Hypersaline sediment
EF688189_Anaerobic wastewater treatment system
EF205555_Geothermal spring mat
EF205517_Geothermal spring mat
FJ638542_Hot spring sediment
AF027081_Hot spring sediment
AF027082_Hot spring sediment
EU635950_Hot spring sediment
JX298789_Hot spring mat
EU924241_Hot spring sediment
AB546037_Oil wellhead
FN436208_Biogas reactor
CU924538_Anaerobic digester
CU918764_Anaerobic digester
AB332115_Methanogenic bioreactor
EF559042_Anaerobic digester
FN436137_Biogas reactor
EF558973_Anaerobic digester
DQ887950_Anaerobic bioreactor
JF417929_Anaerobic digester
DQ887971_Anaerobic bioreactor
FN436086_Biogas reactor
EU812990_Biogas reactor
0.02

Figure 4-1. Phylogenetic tree based on the 16S rRNA gene with sequences from Ca. *Atribacteria* JS1 and OP9 lineages. Bold sequences represent sequences retrieved from single-cell genomes obtained in this study, and c4865 and c14594 were the most abundant Ca. *Atribacteria* JS1 sequences obtained by pyrosequencing in this study. Sequences from other studies are indicated with GenBank accession ID numbers, and the habitats of the organisms from which the sequences were obtained were determined. The lined sequences were obtained from single-cell genomes or metagenomes. Filled circles indicate that the corresponding nodes had more than 60% bootstrap values based on 1,000 resamplings.

The ANI between overlapping regions of the SAGs ranged from 96.8 to 99.9%, with an average of 99.2% per SAG (Appendix 1), and this level was above the ANI cut-off values (95–96%) proposed for delineating bacterial species (Goris et al. 2007, Richter and Rosselló-Móra 2009). Based on the high 16S rRNA gene sequence similarities and ANI values, supporting that the 18 SAGs were from a single species, all data sets were assembled together to construct a cSAG designated RS JS1-cSAG. After the jackknifing procedure to remove chimeric sequences generated during MDA, the assembly size was ~2.24 Mb, with 34.8% GC content (Tables 4-2 and 4-3). The N50 contig size was 5,237 bases, and the largest contig size was 32,879 bases. The number of predicted protein-coding sequences was 2,189, and a total of 831 protein-coding genes were assigned to RAST subsystem categories (Table 4-3). A total of 113 CSCGs were identified using an HMM search against PFAM databases, and the estimated genome completeness of RS JS1-cSAG was found to be 90.32% (Table 4-3 and Appendix 2).

Genome characteristics of RS JS1-cSAG

ABC transporter

ATP-binding cassette (ABC) transporter is a major class of cellular translocation machinery in all bacterial species and often consists of multiple subunits: the ATP-binding protein, the membrane protein, and the substrate-binding protein. RS JS1-cSAG encoded various ABC transporters for amino acids, peptides, oligosaccharides, monosaccharides, minerals, and organic ions (Fig. 4-2). Multiple copies of genes in the RS JS1-cSAG assembly were annotated to ABC transporters of branched-chain amino acids, dipeptides, oligopeptides, ribose, and xylose. Among multiple copies of genes, genes specific to RS

Table 4-3. Assembly statistics and genomic features

General features	RS JS1-cSCG	SL SAG co-assembly
Assembly size (bp)	2,239,755	2,091,128
No. of contigs	694	153
Contig N50, bases	5,237	20,104
Largest contig bases	32,879	64,573
G+C content (%)	34.8	32.48
Protein coding genes	2,189	2,112
Function assigned	831	713
Hypothetical	645	671
rRNA genes	5S(1), 16S(2), and 23S(2)	5S(1), 16S(1), and 23S(1)
tRNA genes	37	22
Genome completeness	90.32% 113/(139*0.9)	81%

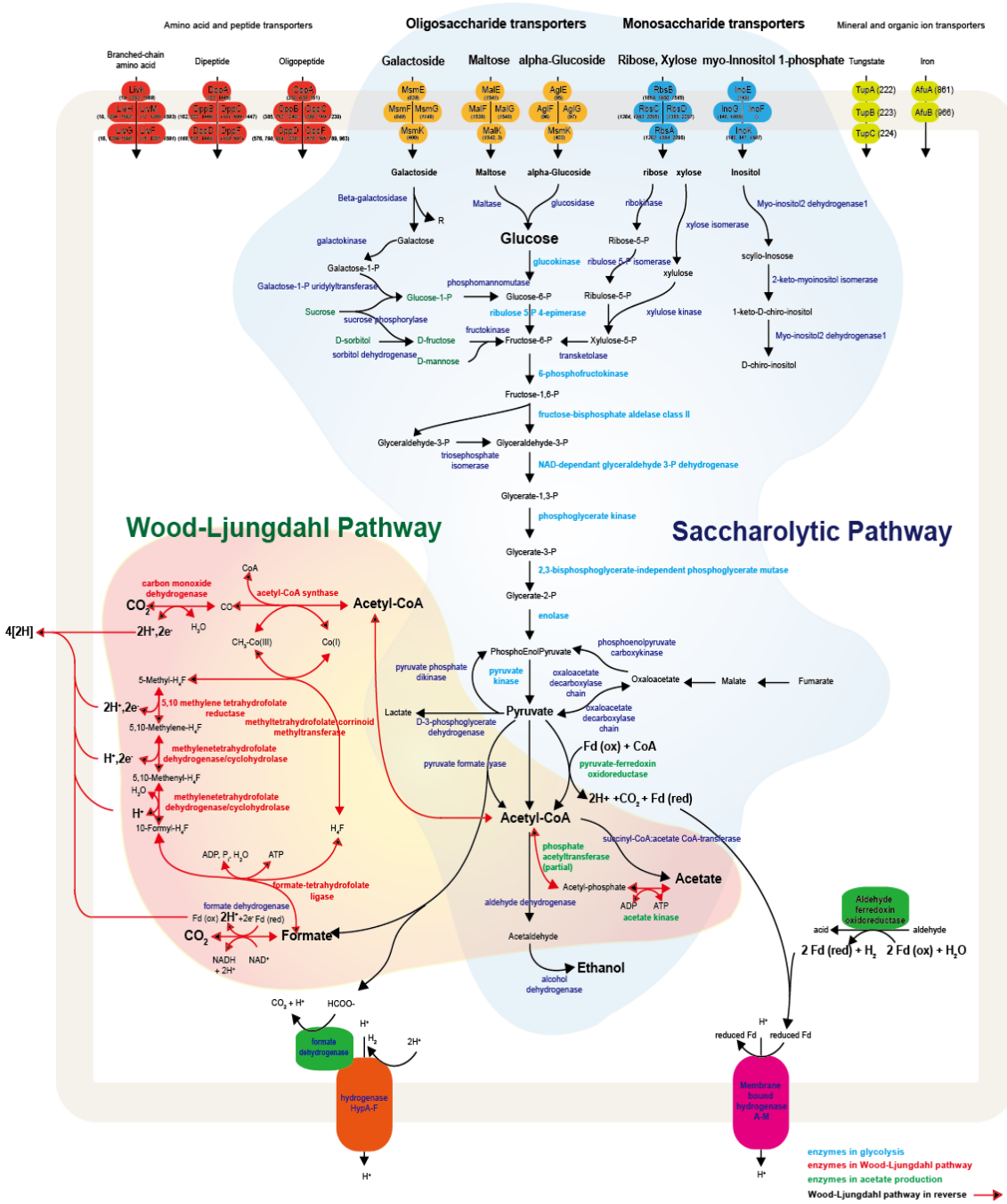


Figure 4-2. Schematic representing metabolic and transport proteins hypothesized from the genome of Ca. *Atribacteria* JS1.

ABC transporters for specific molecules in the bacterial cell membrane are marked by color: amino acid and peptide transporters (red); oligosaccharide transporters (orange); monosaccharide transporters (blue); and mineral and organic ion transporters (green). Gene names and gene numbers in RAST annotations are also indicated. The degradative metabolism of oligosaccharides and monosaccharides and the Wood Ljungdahl pathway were annotated in detail. Proteins included in this schematic of RS JS1-cSAG are listed in Appendix 3. Abbreviations: CoA, coenzyme A; Fd, ferredoxin; Co, cobalamin; H_4F, tetrahydrofolate; ATP, adenosine tri-phosphate; ADP, adenosine diphosphate.

JS1-cSAG with low similarities compared with those reported in other JS1 lineage genomes were found. The ABC transporter gene for maltose and galactoside was found only in RS JS1-cSAG (Fig. 4-2 and Appendix 3).

Central metabolism

RS JS1-cSAG possesses genes involved in the metabolism to convert various sugars transported into the cytosol via ABC transporters to pyruvate. Genes encoding beta-galactosidase, maltase, and glucosidase that cleave the transported oligosaccharides such as galactose, maltose, and alpha-glucoside into monosaccharides were found and a complete set of genes encoding the enzymes of glycolysis, which metabolized the resulting glucose into pyruvate, were identified (Fig. 4-2). RS JS1-cSAG also encoded the enzymes converting other monoscaccharides into intermediates in glycolysis (Fig. 4-2). Genes encoding enzymes catalyzing galactose into glucose-6-phosphate via galactose-1-phosphate, and genes encoding ribokinase and xylose isomerase converting ribose and xylose into xylulose-5-phosphate were present in RS JS1-cSAG (Fig. 4-2). The presence of coding sequences for pyruvate:ferredoxin oxidoreductase suggested the reduction of pyruvate formed from various carbon sources to acetyl-CoA and CO_2 producing reduced ferredoxin. Coding sequences for acetate kinase and alcohol dehydrogenase suggested acetate and ethanol production through the additional fermentation (Fig. 4-2). Reduced ferredoxins can be re-used in the Wood-Ljungdahl pathway, producing acetyl-CoA from CO_2 and CoA. RS JS1-cSAG encoded all components of the Wood-Ljungdahl pathway, such as formate dehydrogenase, formate-tetrahydrofolate ligase, methylenetetrahydrofolate dehydrogenase/cyclohydrolase, 5,10 methylene tetrahydrofolate reductase, acetyl-CoA synthase, pyruvate phosphate dikinase, carbon monoxide dehydrogenase, phosphate acetyltransferase, and acetate kinase (Fig. 4-2).

Comparison of general features with other genomes of *Atribacteria*

Bacterial microcompartments (BMCs) are organelles enveloped with a protein shell that can promote specific metabolic processes by encapsulating and colocalizing enzymes with their substrates and cofactors (Kerfeld et al. 2012). The BMC gene loci in representatives of different lineages of *Atribacteria* were identified (Table 4-4), and BMCs in the *Atribacteria* JS1 lineage were reported to contain aldolases and sugar isomerase, which were not found in BMC loci in other bacterial lineages (Nobu et al. 2015). BMC genes present in other genomes of the *Atribacteria* JS1 lineage were also found in RS JS1-cSAG, suggesting that the BMCs in the *Atribacteria* JS1 lineage may be linked to aldehyde and sugar metabolism.

Another known feature of Ca. *Atribacteria* is the diderm cell envelope structure, which has an outer membrane (Dodsworth et al. 2013b). HMM search results of protein-coding genes against PFAM databases that revealed some proteins, such as BamA protein for outer membrane protein assembly and TolC protein for type I secretin (TolC) involved in diderm cell structure, were present in RS JS1-cSAG (Table 4-5).

Table 4-4. Comparison of BMC gene loci between RS JS1-cSAG and the SL SAG co-assembly

	NCBI contig name in the SL SAG co-assembly	SL SAG co-assembly Protein ID (RAST)	Gene	BMC component	Identities	RS-JS1 cSAG Protein ID (RAST)	RS JS1-cSCG contig no.
BMC	gi|482623737|gb|AQYX01000101.1|	fig|6666666.190053.peg.1785	L-Rhamnose isomerase		92.96	fig|6666666.164181.peg.1515	411
BMC	gi|482623737|gb|AQYX01000101.1|	fig|6666666.190053.peg.1786	Hypothetical protein				
BMC	gi|482623737|gb|AQYX01000101.1|	fig|6666666.190053.peg.1787	Ethanolamine utilization	BMC shell	97.2	fig|6666666.164181.peg.1514	411
BMC	gi|482623737|gb|AQYX01000101.1|	fig|6666666.190053.peg.1788	Ethanolamine utilization	BMC shell	95.83	fig|6666666.164181.peg.1513	411
BMC	gi|482623737|gb|AQYX01000101.1|	fig|6666666.190053.peg.1789	Ethanolamine utilization	BMC shell	91.58	fig|6666666.164181.peg.1512	411
BMC	gi|482623737|gb|AQYX01000101.1|	fig|6666666.190053.peg.1790	Cob(III)alamin reductase		94.52	fig|6666666.164181.peg.1511	411
BMC	gi|482623737|gb|AQYX01000101.1|	fig|6666666.190053.peg.1791	Propanediol utilization	BMC shell	95.73	fig|6666666.164181.peg.1620	433
BMC	gi|482623737|gb|AQYX01000101.1|	fig|6666666.190053.peg.1792	CoA-acylating		91.01	fig|6666666.164181.peg.1642	440
BMC	gi|482623737|gb|AQYX01000101.1|	fig|6666666.190053.peg.1793	Hypothetical protein				
BMC	gi|482623737|gb|AQYX01000101.1|	fig|6666666.190053.peg.1794	Ethanolamine utilization	BMC shell	95.83	fig|6666666.164181.peg.1537	419
BMC	gi|482623737|gb|AQYX01000101.1|	fig|6666666.190053.peg.1795	Ethanolamine utilization	BMC shell	91.86	fig|6666666.164181.peg.1536	419
BMC	gi|482623737|gb|AQYX01000101.1|	fig|6666666.190053.peg.1796	Ribose 5-phosphate		94.54	fig|6666666.164181.peg.1535	419
BMC	gi|482623737|gb|AQYX01000101.1|	fig|6666666.190053.peg.1797	Deoxyribose-phosphate		90.77	fig|6666666.164181.peg.1329	342
BMC	gi|482623737|gb|AQYX01000101.1|	fig|6666666.190053.peg.1798	Ethanolamine utilization				

[1] All assemblies are publicly available at RAST (http://rast.nmpdr.org/) by logging in with username and password “guest”

[2] Re-annotated assemblies for the comparison of RS JS1-cSAG and other *Atribacteria* JS1 lineages in RAST and data are publicly available on RAST.

Table 4-5. Diagnostic markers for the diderm cell envelope structure

			RS JS1-cSAG		SL SAG co-assembly	
Protein name	**Function/annotation**	**Pfam number**	**Protein ID in RAST**	**e-value**	**Protein ID in RAST**	**e-value**
BamA(YaeT)	OM assembly	PF07244	fig|6666666.164181.peg.1041	0.000000015	fig|6666666.138344.peg.944	0.000000047
	OM assembly	PF01103	fig|6666666.164181.peg.1041	2.8E-50	fig|6666666.138344.peg.944	3.20E-50
TolC	Type 1 secretion	PF02321	fig|6666666.164181.peg.1218	0.0000019		
Secretin	Type 2, 3 secretion	PF00263				
Secretin/TonB	Type 2, 3 secretion	PF07660				
FlgH	Flagellar L-ring	PF02107				
FlgI	Flagellar P-ring	PF02119				

4.4 Discussion

The genomes of JS1 obtained in this study showed more than 99.2% of 16S rRNA gene similarity with the predominant JS1 OTU within the core, indicating that the genomes from this study represented the major JS1 species. To date, genomes of the JS1 lineage have been obtained from marine sediments, biofilms from a reactor, and monimolimnion of a meromictic lake (Carr et al. 2015, Dodsworth et al. 2013, Lloyd et al. 2013, Nobu et al. 2015, Rinke et al. 2013). The genomes in our study showed the highest 16S rRNA gene similarity with that of SCGC AD-561-N23 from marine sediments of Adélie Basin, offshore of Antarctica, with a low level of completeness, 15% (Carr et al. 2015). Because the 18 SAGs obtained in this study were considered to be a single species, the genome with the highest coverage and gene count in functional categories among the JS1 genomes, could be obtained by the co-assembly of 18 SAGs. Due to the low coverage of the firstly reported JS1 genome in Antarctic marine sediments, the metabolisms and ecological features were mostly inferred through the comparison with the genomes recovered from other environments and this limited the prediction of *in situ* ecological functions of JS1 there. However, the highest genome of JS1 species which was predominant in the marine sediments of the Ross Sea, made it possible to describe the pathways that explain the fundamental lifestyle in detail and the predominance of this uncultivated group in anaerobic environments.

RS JS1-cSAG showed metabolic potential for acetate formation via fermentation and the Wood-Ljungdahl pathway, containing all components for both pathways (Ljungdahl 1994, Wood et al. 1986). Supporting its nature as a homoacetogen, RS JS1-cSAG was found to harbor ABC transporters for various sugars, such as oligosaccharides and monosaccharides, which can enter the cytosol, and enzymes involved in glycolysis and saccharolytic fermentation. Pyruvate produced from various sugars via glycolysis can then be converted into formate or acetate through various metabolic enzymes, including formate dehydrogenase, pyruvate:ferredoxin oxidoreductase, phosphate acetyltransferase,

acetate kinase, and enzymes in the Wood-Ljungdahl pathway.

In addition, the metabolism inferred from RS JS1-cSAG supported that this bacterium may be metabolically active as a syntrophic acetate-oxidizing bacterium (SAOB). Syntrophic acetate-oxidizing bacteria, such as strain AOR and *Clostridium ultunense* strain BS, are known to grow in syntrophic mode with hydrogenotrophic methanogens by running the Wood-Ljungdahl pathway in reverse (Hattori 2008, Lee and Zinder 1988, Schnürer et al. 1996). Through the syntrophic association, the acetate oxidation, which is thermodynamically unfavorable and cannot proceed in SAOB without the hydrogenotrophic methanogens, becomes possible. Geochemical pore water profiles and methane concentrations were not analyzed in this study. However, seismic evidence for the presence of gas hydrates and free gas in the Victoria Land Basin of the Ross Sea and pockmarks off Franklin Island, indicated the release of subsurface gas near our study site (Geletti and Busetti 2011, Lawver et al. 2012). Thus, RS JS1-cSAG seemed to provide methanogenic substrates, such as acetate from heterotrophic acetogenic substrates, such as galactose, maltose, and ribose via fermentation and Wood-Ljungdahl pathway and hydrogen from syntrophic acetate oxidation of acetate via the reversal of Wood-Ljungdahl pathway, explaining the predominance of the *Atribacteria* JS1 lineage in gas hydrate-bearing or methanogenic marine sediments. Taken together, a hypothetical process for the survival of *Atribacteria* JS1 in anaerobic environments was illustrated in Fig. 4-3.

The highest coverage JS1 genome of the predominant species in the marine sediment of the Ross Sea was obtained. Based on the data presented here, we propose the 'Candidatus *Sediminivita versatilis*' to refer RS JS1-cSAG. The description of the taxa is as follows: '*Sediminivita*' (Se.di.mi.ni.vi'ta. N.L. n. *sedimen* sediment; L. fem. n. *vita* life; N.L. fem. n. *Sediminivita* '*versatilis'* (ver.sa'ti.lis. L. fem. adj. *versatilis* versatile) '*Sediminivita versatilis'* refers to a life isolated from sedimentwhere the lineage is found). 'Candidatus *Sediminivita versatilis*' revealed the metabolic versatility as an acetate producer and syntrophic acetate oxidizer. Although it is

still not clear which lifestyle is predominant in various environments, Candidatus *Sediminivita versatilis* seemed to show increased survivability in nature through syntrophic interactions with partner methanogens or using various heterotrophic substrates to outcompete other primary and secondary fermenters; this versatile growth mode seemed to explain the dominance of Candidatus *Caldatribacterium versatilis* within anoxic sediments of the Ross Sea. Thus, the genomic insights from this study shed light on the ecological function of Ca. *Atribacteria* JS1 in carbon cycling.

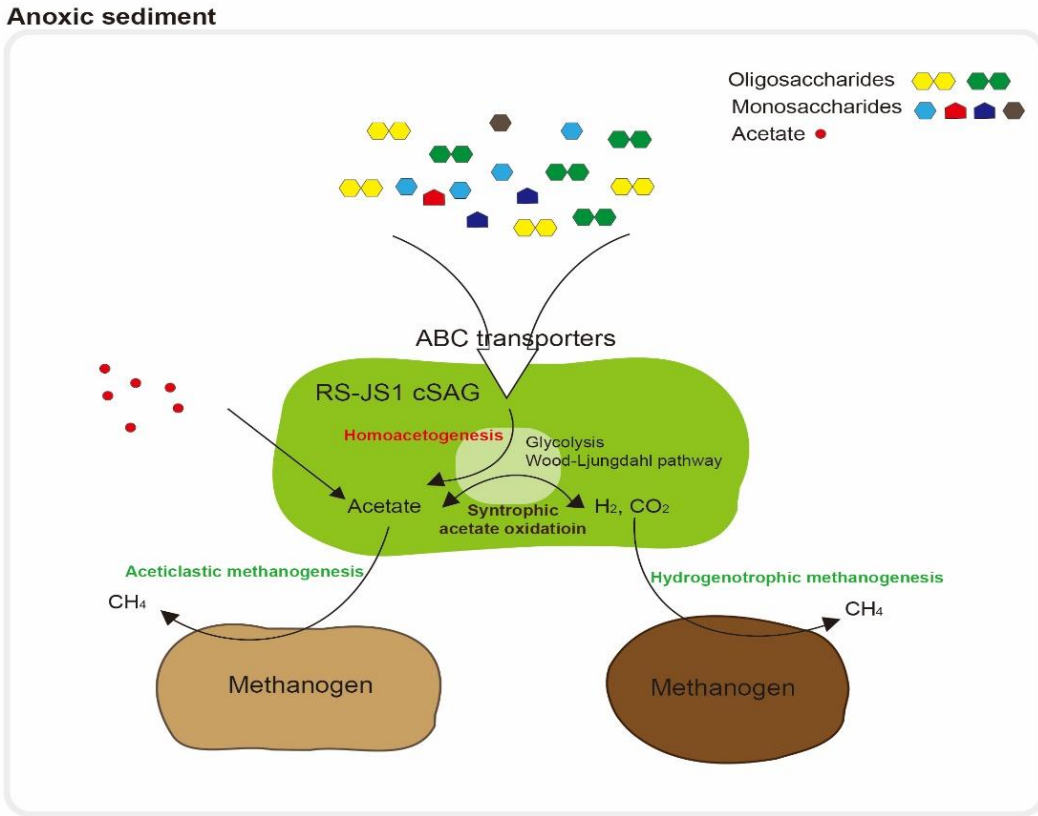


Figure 4-3. Proposed schematic ecological interaction of Ca. *Atribacteria* JS1 with methanogens

CHAPTER 5. Conclusions

In spite of the great expansion of our knowledge on the bacterial diversity with the advancement of sequencing technology, bacterial diversity and ecological functions in SO are still poorly understood. Because of the vulnerability of the SO bacteria to climate change, understanding the diversity and ecological functions of bacterial communities and the environmental factors to influence on them is of significant interest. This study investigated the bacterial diversity at deeper depths and physicochemical parameters that contribute to shape bacterial communities in the SO on regional and local scale. In addition, the physiological characteristics of cultivated bacteria and metabolic potential of uncultivated but dominant group were analyzed.

As a result, highly diverse bacterial groups including previously uncultivated and unidentified groups were revealed and especially, the portion of candidate phylum was higher in the sediments than seawaters. Also, the dominant species in seawaters and sediments were distinct reflecting the physically connected pelagic and the benthic realms differ profoundly in terms of physical, chemical, and biological properties. Distinct distribution pattern according to the circumpolar fronts and depth within a water column, especially oxygen level in the sea water and oxygen availability in the sediments were revealed indicating the oxygen play a constraint for specific bacterial groups.

In spite of the limitation of cultivation, many candidate novel species with less than 98.65% similarity to known strains, were recovered. Some of the culturable bacteria produced cold-active enzymes indicating that they contribute to the hydrolysis of major organic constituents and are therefore involved in carbon and nitrogen cycling at the low temperature of sediments of Ross Sea, Antarctica.

Single cell sorting from the sediment samples with the high portion of Ca. *Atribacteria* JS1 lineage in the anoxic sediments of the Ross Sea and amplification of genomes were performed to understand the ecological function of predominant in the anoxic sediments but uncultivated *Atribacteria* JS1. Successfully, single cell amplified genomes of a single species, which was dominant in the Ross Sea have been obtained. The highest coverage JS1 genome of predominant species in the marine sediment of the Ross Sea revealed how this species is dominant in anaerobic and methanogenic environments. The metabolic versatility of this species as an acetate producer and syntrophic acetate oxidizer seemed to show increased survivability in nature through syntrophic interactions with partner methanogens or using various heterotrophic substrates to outcompete other primary and secondary fermenters; this versatile growth mode seemed to explain the dominance of Ca. *Atribacteria* JS1 within anoxic sediments of the Ross Sea.

Overall, these findings will expand our knowledge about the ecological functions of the marine bacteria in the SO as well as help predict their response to changing environments.

REFERENCES

Abell GCJ, Bowman JP (2005). Ecological and biogeographic relationships of class Flavobacteria in the Southern Ocean. *FEMS Microbiology Ecology* **51:** 265-277.

Albertelli G, Cattaneo-Vietti R, Chiantore M, Pusceddu A, Fabiano M (1998). Food availability to an Adamussium bed during the austral summer 1993/1994 (Terra Nova Bay, Ross Sea). *Journal of Marine Systems* **17:** 425-434.

Amann RI, Ludwig W, Schleifer KH (1995). Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiological Reviews* **59:** 143-169.

Arı´stegui J, Gasol JM, Duarte CM, Herndl GJ (2009). Microbial oceanography of the dark ocean's pelagic realm. *Limnology and Oceanography* **54:** 1501-1529.

Arrigo KR, Robinson DH, Worthen DL, Schieber B, Lizotte MP (1998a). Bio-optical properties of the southwestern Ross Sea. *Journal of Geophysical Research: Oceans* **103:** 21683-21695.

Arrigo KR, Worthen D, Schnell A, Lizotte MP (1998b). Primary production in Southern Ocean waters. *Journal of Geophysical Research: Oceans* **103:** 15587-15600.

Arrigo KR, van Dijken GL (2004). Annual changes in sea-ice, chlorophyll a, and primary production in the Ross Sea, Antarctica. *Deep Sea Research Part II: Topical Studies in Oceanography* **51:** 117-138.

Austen MC, Lambshead PJD, Hutchings PA, Boucher G, Snelgrove PVR, Heip C *et al.* (2002). Biodiversity links above and below the marine sediment–water interface that

may influence community stability. *Biodiversity & Conservation* **11:** 113-136.

Azam F, Fenchel T, Field J, Gray J, Meyerreil L, Thingstad F (1983). The ecological role of water-column microbes in the sea. *Marine Ecology-Progress Series* **10:** 257-263.

Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA *et al.* (2008). The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* **9:** 75.

Bai Y, Yang D, Wang J, Xu S, Wang X, An L (2006). Phylogenetic diversity of culturable bacteria from alpine permafrost in the Tianshan Mountains, northwestern China. *Res Microbiol* **157:** 741-751.

Baldi F, Marchetto D, Pini F, Fani R, Michaud L, Lo Giudice A *et al.* (2010). Biochemical and microbial features of shallow marine sediments along the Terra Nova Bay (Ross Sea, Antarctica). *Continental Shelf Research* **30:** 1614-1625.

Bowman JP, McCammon SA, Gibson JAE, Robertson L, Nichols PD (2003). Prokaryotic metabolic activity and community structure in Antarctic continental shelf sediments. *Applied and Environmental Microbiology* **69:** 2448-2462.

Bowman JP, McCuaig RD (2003). Biodiversity, community structural shifts, and biogeography of prokaryotes within Antarctic continental shelf sediment. *Applied and Environmental Microbiology* **69:** 2463-2483.

Cai W-J, Sayles FL (1996). Oxygen penetration depths and fluxes in marine sediments. *Marine Chemistry* **52:** 123-131.

Carr SA, Vogel SW, Dunbar RB, Brandes J, Spear JR, Levy R *et al.* (2013). Bacterial abundance and composition in marine sediments beneath the Ross Ice Shelf,

Antarctica. *Geobiology* **11:** 377-395.

Carr SA, Orcutt BN, Mandernack KW, Spear JR (2015). Abundant Atribacteria in deep marine sediment from the Adélie Basin, Antarctica. *Frontiers in Microbiology* **6:** 872.

Cavicchioli R (2015). Microbial ecology of Antarctic aquatic systems. *Nature Reviews Microbiology* **13:** 691-706.

Celussi M, Cataletto B, Fonda Umani S, Del Negro P (2009a). Depth profiles of bacterioplankton assemblages and their activities in the Ross Sea. *Deep Sea Research Part I: Oceanographic Research Papers* **56:** 2193-2205.

Celussi M, Paoli A, Crevatin E, Bergamasco A, Margiotta F, Saggiomo V *et al.* (2009b). Short-term under-ice variability of prokaryotic plankton communities in coastal Antarctic waters (Cape Hallett, Ross Sea). *Estuarine, Coastal and Shelf Science* **81:** 491-500.

Celussi M, Bergamasco A, Cataletto B, UmaniI SF, Negro PD (2010). Water masses' bacterial community structure and microbial activities in the Ross Sea, Antarctica. *Antarctic Science* **22:** 361-370.

Chen Y, Murrell JC (2010). When metagenomics meets stable-isotope probing: progress and perspectives. *Trends in Microbiology* **18:** 157-163.

Claire Horner-Devine M, Leibold MA, Smith VH, Bohannan BJM (2003). Bacterial diversity patterns along a gradient of primary productivity. *Ecology Letters* **6:** 613-622.

Clarke K, Gorley R (2006). PRIMER v6: user manual/tutorial (Plymouth routines in

multivariate ecological research). *Plymouth: Primer-E Ltd*.

Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y *et al.* (2013). Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Research*.

Coolen MJL, Overmann J (2000). Functional Exoenzymes as Indicators of Metabolically Active Bacteria in 124,000-Year-Old Sapropel Layers of the Eastern Mediterranean Sea. *Applied and Environmental Microbiology* **66:** 2589-2598.

Corinaldesi C (2015). New perspectives in benthic deep-sea microbial ecology. *Frontiers in Marine Science* **2**.

D/'Hondt S, Inagaki F, Zarikian CA, Abrams LJ, Dubois N, Engelhardt T *et al.* (2015). Presence of oxygen and aerobic communities from sea floor to basement in deep-sea sediments. *Nature Geoscience* **8:** 299-304.

Deacon GER (1982). Physical and biological zonation in the Southern Ocean. *Deep Sea Research Part A Oceanographic Research Papers* **29:** 1-15.

Dodsworth JA, Blainey PC, Murugapiran SK, Swingley WD, Ross CA, Tringe SG *et al.* (2013). Single-cell and metagenomic analyses indicate a fermentative and saccharolytic lifestyle for members of the OP9 lineage. *Nature Communications* **4:** 1854.

Ducklow HW, Erickson M, Kelly J, Montes-Hugo M, Ribic CA, Smith RC *et al.* (2008). Particle export from the upper ocean over the continental shelf of the west Antarctic Peninsula: A long-term record, 1992–2007. *Deep Sea Research Part II: Topical*

Studies in Oceanography **55:** 2118-2131.

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27:** 2194-2200.

Eloe-Fadrosh EA, Ivanova NN, Woyke T, Kyrpides NC (2016). Metagenomics uncovers gaps in amplicon-based detection of microbial diversity. *Nature Microbiology* **1:** 15032.

Fabiano M, Danovaro R, Crisafi E, La Ferla R, Povero P, Acosta-Pomar L (1995). Particulate matter composition and bacterial distribution in Terra Nova Bay (Antarctica) during summer 1989–1990. *Polar Biology* **15:** 393-400.

Fabiano M, Danovaro R (1998). Enzymatic activity, bacterial distribution, and organic matter composition in sediments of the Ross Sea (Antarctica). *Applied and Environmental Microbiology* **64:** 3838-3845.

Fabiano M, Pusceddu A (1998). Total and hydrolizable particulate organic matter (carbohydrates, proteins and lipids) at a coastal station in Terra Nova Bay (Ross Sea, Antarctica). *Polar Biology* **19:** 125-132.

Fabiano M, Danovaro R, Chiantore M, Pusceddu A (2000). Bacteria, protozoa and organic matter composition in the sediments of Terra Nova Bay (Ross Sea). In: Faranda FM, Guglielmo L, Ianora A (eds). *Ross Sea Ecology: Italiantartide Expeditions (1987–1995)*. Springer Berlin Heidelberg: Berlin, Heidelberg. pp 159-169.

Faranda F, Letterio G, Adrianna I (2012). Ross Sea Ecology: Italiantartide Expeditions (1987–1995).

Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL *et al.* (2016). The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Research* **44:** D279-285.

Flombaum P, Gallegos JL, Gordillo RA, Rincón J, Zabala LL, Jiao N *et al.* (2013). Present and future global distributions of the marine Cyanobacteria *Prochlorococcus* and *Synechococcus*. *Proceedings of the National Academy of Sciences* **110:** 9824-9829.

Follows MJ, Dutkiewicz S, Grant S, Chisholm SW (2007). Emergent biogeography of microbial communities in a model ocean. *Science* **315:** 1843-1846.

Fonda Umani S, Monti M, Bergamasco A, Cabrini M, De Vittor C, Burba N *et al.* (2005). Plankton community structure and dynamics versus physical structure from Terra Nova Bay to Ross Ice Shelf (Antarctica). *Journal of Marine Systems* **55:** 31-46.

Freitas S, Hatosy S, Fuhrman JA, Huse SM, Mark Welch DB, Sogin ML *et al.* (2012). Global distribution and diversity of marine Verrucomicrobia. *ISME J* **6:** 1499-1505.

Froneman PW, McQuaid CD, Perissinotto R (1995). Biogeographic structure of the microphytoplankton assemblages of the south Atlantic and Southern Ocean during austral summer. *Journal of Plankton Research* **17:** 1791-1802.

Fuhrman J, Sleeter T, Carlson C, L.M. P (1989). Dominance of bacterial biomass in the Sargasso Sea and its ecological implications. *Marine ecology progress series* **57:** 207-217.

Fuhrman JA, Davis AA (1997). Widespread Archaea and novel Bacteria from the deep sea as shown by 16S rRNA gene sequences. *Marine Ecology Progress Series* **150:** 275-

285.

Fuhrman JA, Steele JA, Hewson I, Schwalbach MS, Brown MV, Green JL *et al.* (2008). A latitudinal diversity gradient in planktonic marine bacteria. *Proceedings of the National Academy of Sciences* **105:** 7774-7778.

Geletti R, Busetti M (2011). A double bottom simulating reflector in the western Ross Sea, Antarctica. *Journal of Geophysical Research* **116**.

Gentile G, Giuliano L, D'Auria G, Smedile F, Azzaro M, De Domenico M *et al.* (2006). Study of bacterial communities in Antarctic coastal waters by a combination of 16S rRNA and 16S rDNA sequencing. *Environmental Microbiology* **8:** 2150-2161.

Gerday C, Aittaleb M, Bentahir M, Chessa JP, Claverie P, Collins T *et al.* (2000). Cold-adapted enzymes: from fundamentals to biotechnology. *Trends in Biotechnolology* **18:** 103-107.

Ghiglione J-F, Galand PE, Pommier T, Pedrós-Alió C, Maas EW, Bakker K *et al.* (2012). Pole-to-pole biogeography of surface and deep marine bacterial communities. *Proceedings of the National Academy of Sciences* **109:** 17633-17638.

Giebel H-A, Brinkhoff T, Zwisler W, Selje N, Simon M (2009). Distribution of Roseobacter RCA and SAR11 lineages and distinct bacterial communities from the subtropics to the Southern Ocean. *Environmental Microbiology* **11:** 2164-2178.

Glud RN (2008). Oxygen dynamics of marine sediments. *Marine Biology Research* **4:** 243-289.

Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM (2007).

DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *International Journal of Systematic and Evolutionary Microbiology* **57:** 81-91.

Groudieva T, Kambourova M, Yusef H, Royter M, Grote R, Trinks H *et al.* (2004). Diversity and cold-active hydrolytic enzymes of culturable bacteria associated with Arctic sea ice, Spitzbergen. *Extremophiles* **8:** 475-488.

Handelsman J (2004). Metagenomics: application of genomics to uncultured microorganisms. *Microbiol Mol Biol Rev* **68:** 669-685.

Hattori S (2008). Syntrophic acetate-oxidizing microbes in methanogenic environments. *Microbes and Environments* **23:** 118-127.

Hauck J, Völker C, Wolf-Gladrow DA, Laufkötter C, Vogt M, Aumont O *et al.* (2015). On the Southern Ocean CO2 uptake and the role of the biological carbon pump in the 21st century *Global Biogeochemical Cycles* **29:** 1451-1470.

Hedlund BP, Dodsworth JA, Murugapiran SK, Rinke C, Woyke T (2014). Impact of single-cell genomics and metagenomics on the emerging view of extremophile "microbial dark matter". *Extremophiles* **18:** 865-875.

Helmke E, Weyland H (2004). Psychrophilic versus psychrotolerant bacteria-occurrence and significance in polar and temperate marine habitats. *Cellular and Molecular Biology* **50:** 553-561.

Herlemann DPR, Labrenz M, Jurgens K, Bertilsson S, Waniek JJ, Andersson AF (2011). Transitions in bacterial communities along the 2000[thinsp]km salinity gradient of the

Baltic Sea. *ISME Journal* **5:** 1571-1579.

Huang S, Wilhelm SW, Harvey HR, Taylor K, Jiao N, Chen F (2012). Novel lineages of Prochlorococcus and Synechococcus in the global oceans. *ISME Journal* **6:** 285-297.

Hug LA, Castelle CJ, Wrighton KC, Thomas BC, Sharon I, Frischkorn KR *et al.* (2013). Community genomic analyses constrain the distribution of metabolic traits across the Chloroflexi phylum and indicate roles in sediment carbon cycling. *Microbiome* **1:** 22.

Humphry DR, George A, Black GW, Cummings SP (2001). *Flavobacterium frigidarium* sp. nov., an aerobic, psychrophilic, xylanolytic and laminarinolytic bacterium from Antarctica. *International Journal of Systematic and Evolutionary Microbiology* **51:** 1235-1243.

Hwang K, Oh J, Kim T-K, Kim BK, Yu DS, Hou BK *et al.* (2013). CLUSTOM: a novel method for clustering 16S rRNA next generation sequences by overlap minimization. *PLOS ONE* **8:** e62623 62621-62612.

Inagaki F, Nunoura T, Nakagawa S, Teske A, Lever M, Lauer A et al. (2006). Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin. *Proceedings of the National Academy of Sciences* **103:** 2815-2820.

Jiang HL, Tay ST, Maszenan AM, Tay JH (2006). Physiological traits of bacterial strains isolated from phenol-degrading aerobic granules. *FEMS Microbiology Ecology* **57:** 182-191.

Kamke J, Sczyrba A, Ivanova N, Schwientek P, Rinke C, Mavromatis K *et al.* (2013).

Single-cell genomics reveals complex carbohydrate degradation patterns in poribacterial symbionts of marine sponges. *ISME Journal* **7:** 2287-2300.

Kato C, Li L, Tamaoka J, Horikoshi K (1997). Molecular analyses of the sediment of the 11000-m deep Mariana Trench. *Extremophiles* **1:** 117-123.

Kerfeld CA, Heinhorst S, Cannon GC (2012). Bacterial Microcompartments. *Annual Review of Microbiology* **64:** 391-408.

Kim M, Oh H-S, Park S-C, Chun J (2014). Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *International Journal of Systematic and Evolutionary Microbiology* **64:** 346-351.

Kim O-S, Cho Y-J, Lee K, Yoon S-H, Kim M, Na H *et al.* (2012). Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. *International Journal of Systematic and Evolutionary Microbiology* **62:** 716-721.

Kim SJ, Yim JH (2007). Cryoprotective properties of exopolysaccharide (P-21653) produced by the Antarctic bacterium, *Pseudoalteromonas arctica* KOPRI 21653. *Journal of Microbiol*ogy **45:** 510-514.

Kimura M (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Moecularl Evolution* **16:** 111-120.

Kirchman DL (2002). The ecology of Cytophaga–Flavobacteria in aquatic environments.

FEMS Microbiology Ecology **39:** 91-100.

Kirchman DL, Moran XAG, Ducklow H (2009). Microbial growth in the polar oceans [mdash] role of temperature and potential impact of climate change. *Nature Reviews Microbiology* **7:** 451-459.

Kolinko S, Jogler C, Katzmann E, Wanner G, Peplies J, Schüler D (2012). Single-cell analysis reveals a novel uncultivated magnetotactic bacterium within the candidate division OP3. *Environmental Microbiology* **14:** 1709-1721.

Kouridaki I, Polymenakou PN, Tselepides A, Mandalakis M, Kenneth L. Smith J (2010). Phylogenetic diversity of sediment bacteria from the deep Northeastern Pacific Ocean: a comparison with the deep Eastern Mediterranean Sea. *International Microbiology* **13:** 143-150.

Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C *et al.* (2004). Versatile and open software for comparing large genomes. *Genome Biology* **5:** R12.

López-García P, López-López A, Moreira D, Rodríguez-Valera F (2001). Diversity of free-living prokaryotes from a deep-sea site at the Antarctic Polar Front. *FEMS Microbiology Ecology* **36:** 193-202.

Lane (1991). 16S/23S rRNA sequencing In E. Stackebrandt and M. Goodfellow (ed.), Nucleic acid techniques in bacterial systematics. *John Wiley and Sons, New York, NY***:** 115-175.

Lawver L, Lee J, Kim Y, Davey F (2012). Flat-topped mounds in western Ross Sea: Carbonate mounds or subglacial volcanic features? *Geosphere* **8:** 645-653.

Learman DR, Henson MW, Thrash JC, Temperton B, Brannock PM, Santos SR *et al.* (2016). Biogeochemical and microbial variation across 5500 km of Antarctic surface sediment implicates organic matter as a driver of benthic community structure. *Frontiers in Microbiology* **7:** 284.

Lee MJ, Zinder SH (1988). Isolation and characterization of a thermophilic bacterium which oxidizes acetate in syntrophic association with a methanogen and which grows acetogenically on H2-CO2. *Applied and Environmental Microbiology* **54:** 124-129.

Lee YM, Kim G, Jung Y-J, Choe C-D, Yim JH, Lee HK et al. (2012). Polar and Alpine Microbial Collection (PAMC): a culture collection dedicated to polar and alpine microorganisms. *Polar Biology* **35:** 1433-1438.

Lee YM, Jung Y-J, Hong SG, Kim JH, Lee HK (2014). Diversity and physiological characteristics of culturable bacteria from marine sediments of Ross Sea, Antarctica. *Korean Journal of Microbiology* **50:** 119-127.

Li H, Yu Y, Luo W, Zeng Y, Chen B (2009). Bacterial diversity in surface sediments from the Pacific Arctic Ocean. *Extremophiles* **13:** 233-246.

Ljungdahl LG (1994). The Acetyl-CoA Pathway and the Chemiosmotic Generation of ATP during Acetogenesis. In: Drake HL (ed). *Acetogenesis*. Springer US: Boston, MA. pp 63-87.

Lloyd KG, Schreiber L, Petersen DG, Kjeldsen KU, Lever MA, Steen AD *et al.* (2013). Predominant archaea in marine sediments degrade detrital proteins. *Nature* **496:** 215-218.

Lo Giudice A, Caruso C, Mangano S, Bruni V, De Domenico M, Michaud L (2012). Marine bacterioplankton diversity and community composition in an Antarctic Coastal Environment. *Microbial Ecology* **63:** 210-223.

Logares R, Lindstrom ES, Langenheder S, Logue JB, Paterson H, Laybourn-Parry J *et al.* (2013). Biogeography of bacterial communities exposed to progressive long-term environmental change. *ISME Journal* **7:** 937-948.

Lozupone CA, Knight R (2007). Global patterns in bacterial diversity. *Proceedings of the National Academy of Sciences* **104:** 11436-11440.

Macaulay IC, Voet T (2014). Single cell genomics: advances and future perspectives. *PLoS Genetics* **10:** e1004126.

Macura D, Townsley PM (1984). Scandinavian Ropy Milk — Identification and Characterization of Endogenous Ropy Lactic Streptococci and Their Extracellular Excretion1. *Journal of Dairy Science* **67:** 735-744.

Mangoni O, Modigh M, Conversano F, Carrada GC, Saggiomo V (2004). Effects of summer ice coverage on phytoplankton assemblages in the Ross Sea, Antarctica. *Deep Sea Research Part I: Oceanographic Research Papers* **51:** 1601-1617.

Marshall IP, Blainey PC, Spormann AM, Quake SR (2012). A single-cell genome for Thiovulum sp. *Applied and Environmental Microbiology* **78:** 8555-8563.

McCammon SA, Innes BH, Bowman JP, Franzmann PD, Dobson SJ, Holloway PE *et al.* (1998). *Flavobacterium hibernum* sp. nov., a lactose-utilizing bacterium from a freshwater Antarctic Lake. *International Journal of Systematic Bacteriology* **48:**

1405-1412.

McCammon SA, Bowman JP (2000). Taxonomy of Antarctic *Flavobacterium* species: description of *Flavobacterium gillisiae* sp. nov., *Flavobacterium tegetincola* sp. nov., and *Flavobacterium xanthum* sp. nov., nom. rev. and reclassification of [*Flavobacterium*] *salegens* as *Salegentibacter salegens* gen. nov., comb. nov. *International Journal of Systematic and Evolutionary Microbiology* **50:** 1055-1063.

Michaud L, Cello DF, Brilli M, Fani R, Giudice LA, Bruni V (2004). Biodiversity of cultivable psychrotrophic marine bacteria isolated from Terra Nova Bay (Ross Sea, Antarctica). *FEMS Microbiology Letters* **230:** 63-71.

Michaud L, Caruso C, Mangano S, Interdonato F, Bruni V, Lo Giudice A (2012). Predominance of *Flavobacterium*, *Pseudomonas*, and *Polaromonas* within the prokaryotic community of freshwater shallow lakes in the northern Victoria Land, East Antarctica. *FEMS Microbiology Ecology* **82:** 391-404.

Mikaloff Fletcher SE, Gruber N, Jacobson AR, Doney SC, Dutkiewicz S, Gerber M et al. (2006). Inverse estimates of anthropogenic CO2 uptake, transport, and storage by the ocean. *Global Biogeochemical Cycles* **20.**

Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M (2007). KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Research* **35:** 182-185.

Murray AE, Grzymski JJ (2007). Diversity and genomics of Antarctic marine micro-organisms. *Philosophical Transactions of the Royal Society B: Biological Sciences* **362:** 2259-2271.

Nobu MK, Dodsworth JA, Murugapiran SK, Rinke C, Gies EA, Webster G *et al.* (2015a). Phylogeny and physiology of candidate phylum /`Atribacteria/' (OP9/JS1) inferred from cultivation-independent genomics. *ISME Journal* **10:** 273-286.

Nobu MK, Narihiro T, Rinke C, Kamagata Y, Tringe SG, Woyke T *et al.* (2015b). Microbial dark matter ecogenomics reveals complex synergistic networks in a methanogenic bioreactor. *ISME Journal* **9:** 1710-1722.

Nunoura T, Takaki Y, Shimamura S, Kakuta J, Kazama H, Hirai M *et al.* (2016). Variance and potential niche separation of microbial communities in subseafloor sediments off Shimokita Peninsula, Japan. *Environmental Microbiology* **18:** 1889-1906.

Oh J, Kim B, Cho W-S, Hong S, Kim K (2012). PyroTrimmer: a software with GUI for pre-processing 454 amplicon sequences. *Journal of Microbiology* **50:** 766-769.

Orcutt BN, Sylvan JB, Knab NJ, Edwards KJ (2011). Microbial ecology of the dark ccean above, at, and below the seafloor. *Microbiology and Molecular Biology Reviews* **75:** 361-422.

Orsi AH, Whitworth Iii T, Nowlin Jr WD (1995a). On the meridional extent and fronts of the Antarctic Circumpolar Current. *Deep Sea Research Part I: Oceanographic Research Papers* **42:** 641-673.

Pakhomov AE, McQuaid DC (1996). Distribution of surface zooplankton and seabirds across the Southern Ocean. *Polar Biology* **16:** 271-286.

Pakhomov EA, Perissinotto R, Froneman PW (1999). Predation impact of carnivorous macrozooplankton and micronekton in the Atlantic sector of the Southern Ocean.

Journal of Marine Systems **19:** 47-64.

Pakhomov EA, Perissinotto R, McQuaid CD, Froneman PW (2000). Zooplankton structure and grazing in the Atlantic sector of the Southern Ocean in late austral summer 1993: Part 1. Ecological zonation. *Deep Sea Research Part I: Oceanographic Research Papers* **47:** 1663-1686.

Park S-J, Park B-J, Jung M-Y, Kim S-J, Chae J-C, Roh Y et al. (2011). Influence of Deglaciation on Microbial Communities in Marine Sediments Off the Coast of Svalbard, Arctic Circle. *Microbial Ecology* **62:** 537-548.

Parkes RJ, Cragg B, Roussel E, Webster G, Weightman A, Sass H (2014). A review of prokaryotic populations and processes in sub-seafloor sediments, including biosphere:geosphere interactions. *Marine Geology* **352:** 409-425.

Partensky F, Hess WR, Vaulot D (1999). Prochlorococcus, a Marine Photosynthetic Prokaryote of Global Significance. *Microbiology and Molecular Biology Reviews* **63:** 106-127.

Polymenakou P, Bertilsson S, Tselepides A, Stephanou E (2005). Bacterial community composition in different sediments from the eastern Mediterranean Sea: a comparison of four 16S ribosomal DNA clone libraries. *Microbial Ecology* **50:** 447-462.

Polymenakou PN, Lampadariou N, Mandalakis M, Tselepides A (2009). Phylogenetic diversity of sediment bacteria from the southern Cretan margin, Eastern Mediterranean Sea. *Systematic and Applied Microbiology* **32:** 17-26.

Povero P, Castellano M, Ruggieri N, Monticelli LS, Saggiomo V, Chiantore M *et al.* (2006).

Water column features and their relationship with sediments and benthic communities along the Victoria Land coast, Ross Sea, summer 2004. *Antarctic Science* **18:** 603-613.

Pusceddu A, Dell'Anno A, Fabiano M (2000). Organic matter composition in coastal sediments at Terra Nova Bay (Ross Sea) during summer 1995. *Polar Biology* **23:** 288-293.

Qian P-Y, Wang Y, Lee OO, Lau SCK, Yang J, Lafi FF *et al.* (2011). Vertical stratification of microbial communities in the Red Sea revealed by 16S rDNA pyrosequencing. *ISME Journal* **5:** 507-518.

Radajewski S, Ineson P, Parekh NR, Murrell JC (2000). Stable-isotope probing as a tool in microbial ecology. *Nature* **403:** 646-649.

Ravenschlag K, Sahm K, Amann R (2001). Quantitative molecular analysis of the microbial community in marine Arctic sediments (Svalbard). *Applied and Environmental Microbiology* **67:** 387-395.

Revsbech NP, Jørgensen BB, Blackburn TH (1980). Oxygen in the Sea Bottom Measured with a Microelectrode. *Science* **207:** 1355-1356.

Richter M, Rosselló-Móra R (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proceedings of the National Academy of Sciences* **106:** 19126-19131.

Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng J-F *et al.* (2013a). Insights into the phylogeny and coding potential of microbial dark matter. *Nature* **499:**

431-437.

Russell N (1998). Molecular adaptations in psychrophilic bacteria: Potential for biotechnological applications. In: Antranikian G (ed). *Biotechnology of Extremophiles*. Springer Berlin Heidelberg. pp 1-21.

Saitou N, Nei M (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4:** 406-425.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB *et al.* (2009). Introducing mothur: open-source, platform-Independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology* **75:** 7537-7541.

Schnürer A, Schink B, Svensson BH (1996). Clostridium ultunense sp. nov., a mesophilic bacterium oxidizing acetate in syntrophic association with a hydrogenotrophic methanogenic bacterium. *International Journal of Systematic and Evolutionary Microbiology* **46:** 1145-1152.

Selbmann L, Onofri S, Fenice M, Federici F, Petruccioli M (2002). Production and structural characterization of the exopolysaccharide of the Antarctic fungus *Phoma herbarum* CCFEE 5080. *Research in Microbiolology* **153:** 585-592.

Selje N, Simon M, Brinkhoff T (2004). A newly discovered Roseobacter cluster in temperate and polar oceans. *Nature* **427:** 445-448.

Ship S, Anderson J, Domack E (1999). Late Pleistocene–Holocene retreat of the West Antarctic Ice-Sheet system in the Ross Sea: Part 1—Geophysical results. *Geological*

Society of America Bulletin **111:** 1486-1516.

Signori CN, Thomas F, Enrich-Prast A, Pollery RCG, Sievert SM (2014). Microbial diversity and community structure across environmental gradients in Bransfield Strait, Western Antarctic Peninsula. *Frontiers in Microbiology* **5**.

Sokolov S, Rintoul SR (2007). On the relationship between fronts of the Antarctic Circumpolar Current and surface chlorophyll concentrations in the Southern Ocean. *Journal of Geophysical Research: Oceans* **112**.

Staley JT, Herwig RP (1993). Degradation of particulate organic material in the Antarctic. *In EI Friedmann (ed), Antarctic Microbiology Wiley-Liss Inc, New York, NY, USA***:** 241-264.

Sunagawa S, Coelho LP, Chaffron S, Kultima JR, Labadie K, Salazar G *et al.* (2015). Structure and function of the global ocean microbiome. *Science* **348**.

Talley LD, L.Pickard G, J.Emery W, H.Swift J (2011). Descriptive physical ocenography: an introduction**:** 437-471.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013). MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* **30:** 2725-2729.

Thomas MK, Kremer CT, Klausmeier CA, Litchman E (2012). A global pattern of thermal adaptation in marine phytoplankton. *Science* **338:** 1085-1088.

Van Trappen S, Mergaert J, Swings J (2003). *Flavobacterium gelidilacus* sp. nov., isolated from microbial mats in Antarctic lakes. *International Journal of Systematic and*

Evolutionary Microbiology **53:** 1241-1245.

Van Trappen S, Vandecandelaere I, Mergaert J, Swings J (2004). *Flavobacterium degerlachei* sp. nov., Flavobacterium frigoris sp. nov. and Flavobacterium micromati sp. nov., novel psychrophilic bacteria isolated from microbial mats in Antarctic lakes. *International Journal of Systematic and Evolutionary Microbiology* **54:** 85-92.

Van Trappen S, Vandecandelaere I, Mergaert J, Swings J (2005). *Flavobacterium fryxellicola* sp. nov. and Flavobacterium psychrolimnae sp. nov., novel psychrophilic bacteria isolated from microbial mats in Antarctic lakes. *International Journal of Systematic and Evolutionary Microbiology* **55:** 769-772.

Vazquez SC, Coria SH, Mac Cormack WP (2004). Extracellular proteases from eight psychrotolerant Antarctic strains. *Microbiological Research* **159:** 157-166.

Walker TR (2005). Distribution of oxygen, sulfides and optimum temperature for sulfate reduction in Antarctic marine sediments. *Polish Polar Research* **26:** 215-230.

Ward P, Whitehouse M, Brandon M, Shreeve R, Woodd-Walker R (2003). Mesozooplankton community structure across the Antarctic Circumpolar Current to the north of South Georgia: Southern Ocean. *Marine Biology* **143:** 121-130.

Webster G, John Parkes R, Cragg BA, Newberry CJ, Weightman AJ, Fry JC (2006a). Prokaryotic community composition and biogeochemical processes in deep subseafloor sediments from the Peru Margin. *FEMS Microbiology Ecology* **58:** 65-85.

Webster G, Watt LC, Rinna J, Fry JC, Evershed RP, Parkes RJ *et al.* (2006b). A comparison of stable-isotope probing of DNA and phospholipid fatty acids to study prokaryotic

functional diversity in sulfate-reducing marine sediment enrichment slurries. *Environmental Microbiology* **8:** 1575-1589.

Webster G, Sass H, Cragg BA, Gorra R, Knab NJ, Green CJ *et al.* (2011). Enrichment and cultivation of prokaryotes associated with the sulphate–methane transition zone of diffusion-controlled sediments of Aarhus Bay, Denmark, under heterotrophic conditions. *FEMS Microbiology Ecology* **77:** 248-263.

Whitman WB, Coleman DC, Wiebe WJ (1998). Prokaryotes: the unseen majority. *Proceedings of the National Academy of Sciences* **95:** 6578-6583.

Wilkins D, Lauro FM, Williams TJ, Demaere MZ, Brown MV, Hoffman JM *et al.* (2012). Biogeographic partitioning of Southern Ocean microorganisms revealed by metagenomics. *Environmental Microbiology* **15:** 1318-1333.

Wilkins D, van Sebille E, Rintoul SR, Lauro FM, Cavicchioli R (2013). Advection shapes Southern Ocean microbial assemblages independent of distance and environment effects. *Nature Communications* **4**.

Williams TJ, Wilkins D, Long E, Evans F, DeMaere MZ, Raftery MJ *et al.* (2013). The role of planktonic Flavobacteria in processing algal organic matter in coastal East Antarctica revealed using metagenomics and metaproteomics. *Environmental Microbiology* **15:** 1302-1317.

Wood HG, Ragsdale SW, Pezacka E (1986). The acetyl-CoA pathway: a newly discovered pathway of autotrophic growth. *Trends in Biochemical Sciences* **11:** 14-18.

Yakimov MM, Gentile G, Bruni V, Cappello S, D'Auria G, Golyshin PN *et al.* (2004).

Crude oil-induced structural shift of coastal bacterial communities of rod bay (Terra Nova Bay, Ross Sea, Antarctica) and characterization of cultured cold-adapted hydrocarbonoclastic bacteria. *FEMS Microbiology Ecology* **49:** 419-432.

Yi H, Oh H-M, Lee J-H, Kim S-J, Chun J (2005). *Flavobacterium antarcticum* sp. nov., a novel psychrotolerant bacterium isolated from the Antarctic. *International Journal of Systematic and Evolutionary Microbiology* **55:** 637-641.

Yi H, Chun J (2006). *Flavobacterium weaverense* sp. nov. and *Flavobacterium segetis* sp. nov., novel psychrophiles isolated from the Antarctic. *International Journal of Systematic and Evolutionary Microbiology* **56:** 1239-1244.

Yu Y, Li HR, Zeng YX, Chen B (2011). Bacterial diversity and bioprospecting for cold-active hydrolytic enzymes from culturable bacteria associated with sediment from Nella Fjord, Eastern Antarctica. *Marine Drugs* **9:** 184-195.

Zinger L, Amaral-Zettler LA, Fuhrman JA, Horner-Devine MC, Huse SM, Welch DBM *et al.* (2011). Global patterns of bacterial beta-diversity in seafloor and seawater ecosystems. *PLoS ONE* **6:** e24570.

APPENDICES

Appendix 1. Average nucleotide identity comparisons

	AG18 3A05	AG18 3A11	AG18 3B19	AG18 3C13	AG18 3D10	AG18 3D19	AG18 3F16	AG18 3G20	AG1 83J17	AG18 3M21	AG18 3N04	AG18 3N14	AG18 3O15	AG18 3P03	AG18 3P08	AG18 3P13	AG18 3P14	AG18 3P18
AG18		99.48	99.78	99.48	99.62	99.71	99.57	99.82	99.4	99.43	99.56	99.56	99.66	98.64	99.74	99.42	99.85	98.2
AG18	99.48		99.46	99.31	99.05	99.5	99.36	99.6	99.4	99.41	99.18	99.24	99.47	98.02	99.56	99.38	99.44	**97.77**
AG18	99.77	99.46		99.32	99.55	99.73	99.52	99.81	99.53	99.64	99.49	99.31	99.59	**97.64**	99.71	99.37	99.74	**97.99**
AG18	99.48	99.31	99.32		99.04	99.48	99.27	99.6	99.2	99.3	99.23	99.23	99.48	**97.88**	99.53	99.22	99.38	**96.77**
AG18	99.62	99.05	99.55	99.1		99.72	99.25	99.57	98.24	99.37	99.34	99.3	99.26	98.14	99.7	99.51	99.48	98.17
AG18	99.71	99.5	99.73	99.48	99.72		99.24	99.81	99.31	99.7	99.42	99.47	99.6	98.53	99.72	99.78	99.84	98.31
AG18	99.59	99.36	99.52	99.27	99.26	99.24		99.66	99.18	99.44	99.43	99.33	99.55	**97.66**	99.65	99.3	99.58	**97.15**
AG18	99.82	99.6	99.81	99.6	99.57	99.81	99.66		99.51	99.51	99.46	99.49	99.63	98.46	99.84	99.62	99.77	98
AG18	99.4	99.4	99.53	99.2	98.39	99.31	99.18	99.51		99.39	98.96	98.99	99.38	98.58	99.48	99.46	99.6	98.42
AG18	99.43	99.41	99.64	99.3	99.37	99.7	99.44	99.51	99.38		99.33	99.35	99.47	**97.77**	99.65	99.32	99.33	**97.57**
AG18	99.56	99.18	99.49	99.23	99.34	99.42	99.43	99.46	98.98	99.31		99.78	99.39	98.15	99.64	99.18	99.49	**97.88**

AG18	99.56	99.22	99.31	99.23	99.32	99.47	99.33	99.48	98.96	99.35	99.78		99.38	98.21	99.6	99.15	99.49	**97.87**
AG18	99.66	99.47	99.59	99.49	99.25	99.6	99.56	99.64	99.36	99.45	99.37	99.37		**97.84**	99.54	99.56	99.6	**97.75**
AG18	98.64	98.02	**97.66**	**97.88**	98.14	98.53	**97.68**	98.46	98.58	**97.72**	98.17	98.23	**97.86**		98.77	98.42	98.47	98.71
AG18	99.76	99.56	99.71	99.53	99.7	99.72	99.65	99.84	99.48	99.65	99.64	99.6	99.54	98.75		99.66	99.79	98.48
AG18	99.42	99.38	99.37	99.22	99.51	99.78	99.3	99.62	99.46	99.29	99.15	99.14	99.56	98.39	99.66		99.31	**97.36**
AG18	99.87	99.44	99.74	99.38	99.48	99.84	99.58	99.77	99.6	99.33	99.49	99.49	99.6	98.47	99.79	99.31		**97.64**
AG18	98.2	**97.77**	**97.99**	**96.77**	98.17	98.32	**97.15**	98	98.35	**97.56**	**97.88**	**97.87**	**97.75**	98.71	98.48	**97.36**	**97.68**	

Appendix 2. Conserved single copy gene sets

	PFAM ID	HMM name	HMM cutoff	Description of bacterial genes	RS-JS1-cSAG Protein ID in RAST
1	PF03485	Arg_tRNA_synt_N	32.55	Arginyl tRNA synthetase N terminal domain	
2	PF03484	B5	26.55	tRNA synthetase B5 domain	fig|6666666.164181.peg.1887
3	PF01121	CoaE	84.65	Dephospho-CoA kinase	fig|6666666.164181.peg.2082
4	PF03772	Competence	82.7	Competence protein	fig|6666666.164181.peg.1381
5	PF03602	Cons_hypoth95	38.8	Conserved hypothetical protein 95	fig|6666666.164181.peg.1849
6	PF06418	CTP_synth_N	208	CTP synthase N-terminus	fig|6666666.164181.peg.1206
7	PF02224	Cytidylate_kin	79.35	Cytidylate kinase	fig|6666666.164181.peg.1846
8	PF00712	DNA_pol3_beta	48.7	DNA polymerase III beta subunit, N-terminal domain	
9	PF02767	DNA_pol3_beta_2	49.65	DNA polymerase III beta subunit, central domain	
10	PF02768	DNA_pol3_beta_3	44.5	DNA polymerase III beta subunit, C-terminal domain	
11	PF00035	dsrm	20.2	Double-stranded RNA binding motif	fig|6666666.164181.peg.1918
12	PF00889	EF_TS	108.35	Elongation factor TS	fig|6666666.164181.peg.1923
13	PF01176	eIF-1a	39.25	Translation initiation factor 1A / IF-1	fig|6666666.164181.peg.319
14	PF00113	Enolase_C	206.35	Enolase, C-terminal TIM barrel domain	fig|6666666.164181.peg.336
15	PF03952	Enolase_N	92.45	45 Enolase, N-terminal domain	fig|6666666.164181.peg.336
16	PF06574	FAD_syn	73.25	25 FAD synthetase	
17	PF03147	FDX-ACB	39.75	Ferredoxin-fold anticodon binding domain	fig|6666666.164181.peg.1887
18	PF01687	Flavokinase	59.3	Riboflavin kinase	

Appendix 2. To be continued

	PFAM ID	HMM name	HMM cutoff	Description of bacterial genes	RS-JS1-cSAG Protein ID in RAST
19	PF02938	GAD	40.05	GAD domain	fig|6666666.164181.peg.918
20	PF02527	GidB	77.85	rRNA small subunit methyltransferase G	fig|6666666.164181.peg.732
21	PF00958	GMP_synt_C	61.3	GMP synthase C terminal domain	
22	PF01025	GrpE	72.25	GrpE	fig|6666666.164181.peg.1237
23	PF01018	GTP1_OBG	98.9	GTP1/OBG	fig|6666666.164181.peg.352
24	PF11987	IF?2	61.7	Translation-initiation factor 2	fig|6666666.164181.peg.771
25	PF04760	IF2_N	29.95	Translation initiation factor IF-2, N-terminal region	fig|6666666.164181.peg.1943
26	PF00707	IF3_C	51.55	Translation initiation factor IF-3, C-terminal domain	fig|6666666.164181.peg.49
27	PF05198	IF3_N	47.95	Translation initiation factor IF-3, N-terminal domain	
28	PF01715	IPPT	129.15	IPP transferase	fig|6666666.164181.peg.721
29	PF06421	LepA_C	79.7	GTP-binding protein LepA C-terminus	fig|6666666.164181.peg.1373
30	PF01795	Methyltransf_5	174.05	MraW methylase family	
31	PF02873	MurB_C	46.15	UDP-N-acetylenolpyruvoylglucosamine reductase, C-terminal domain	fig|6666666.164181.peg.2
32	PF08529	NusA_N	61.55	NusA N-terminal domain	fig|6666666.164181.peg.1941
33	PF02410	Oligomerisation	44.9	Oligomerisation domain	fig|6666666.164181.peg.731
34	PF01195	Pept_tRNA_hydro	99.15	Peptidyl-tRNA hydrolase	fig|6666666.164181.peg.1712
35	PF01252	Peptidase_A8	61.35	Signal peptidase (SPase) II	fig|6666666.164181.peg.8
36	PF00162	PGK	236	Phosphoglycerate kinase	fig|6666666.164181.peg.1054
37	PF02912	Phe_tRNA-synt_N	33.55	Aminoacyl tRNA synthetase class II, N-terminal domain	fig|6666666.164181.peg.181
38	PF03726	PNPase	25.6	Polyribonucleotide nucleotidyltransferase, RNA binding domain	fig|6666666.164181.peg.1716
39	PF01416	PseudoU_synth_1	42.45	tRNA pseudouridine synthase	fig|6666666.164181.peg.1615
40	PF02033	RBFA	44.9	Ribosome-binding factor A	fig|6666666.164181.peg.769
41	PF00154	RecA	276.55	recA bacterial DNA recombination protein	fig|6666666.164181.peg.1191

Appendix 2. To be continued

	PFAM ID	HMM name	HMM cutoff	Description of bacterial genes	RS-JS1-cSAG Protein ID in RAST
42	PF02132	RecR	21.25	RecR protein	fig|6666666.164181.peg.1030
43	PF00825	Ribonuclease_P	40.55	Ribonuclease P	
44	PF00687	Ribosomal_L1	75.1	Ribosomal protein L1p/L10e family	fig|6666666.164181.peg.646
45	PF00466	Ribosomal_L10	37.3	Ribosomal protein L10	fig|6666666.164181.peg.645
46	PF00298	Ribosomal_L11	42.65	Ribosomal protein L11, RNA binding domain	fig|6666666.164181.peg.647
47	PF03946	Ribosomal_L11_N	45.9	Ribosomal protein L11, N-terminal domain	fig|6666666.164181.peg.647
48	PF00542	Ribosomal_L12	44	Ribosomal protein L7/L12 C-terminal domain	fig|6666666.164181.peg.644
49	PF00572	Ribosomal_L13	81.3	Ribosomal protein L13	fig|6666666.164181.peg.1853
50	PF00238	Ribosomal_L14	79.55	Ribosomal protein L14p/L23e	fig|6666666.164181.peg.1704
51	PF00252	Ribosomal_L16	77.45	Ribosomal protein L16p/L10e	fig|6666666.164181.peg.1707
52	PF01196	Ribosomal_L17	55.75	Ribosomal protein L17	fig|6666666.164181.peg.325
53	PF00828	Ribosomal_L18e	37.55	Ribosomal protein L18e/L15	fig|6666666.164181.peg.315
54	PF00861	Ribosomal_L18p	54.85	Ribosomal L18p/L5e family	fig|6666666.164181.peg.1071
55	PF01245	Ribosomal_L19	71.7	Ribosomal protein L19	
56	PF00181	Ribosomal_L2	52.45	Ribosomal Proteins L2, RNA binding domain	fig|6666666.164181.peg.1047
57	PF03947	Ribosomal_L2_C	87.3	Ribosomal Proteins L2, C-terminal domain	fig|6666666.164181.peg.1047
58	PF00453	Ribosomal_L20	70.5	Ribosomal protein L20	fig|6666666.164181.peg.47
59	PF00829	Ribosomal_L21p	53.45	Ribosomal prokaryotic L21 protein	fig|6666666.164181.peg.350
60	PF00237	Ribosomal_L22	57.6	Ribosomal protein L22p/L17e	fig|6666666.164181.peg.1049
61	PF00276	Ribosomal_L23	37.45	Ribosomal protein L23	fig|6666666.164181.peg.1046
62	PF01016	Ribosomal_L27	55.4	Ribosomal L27 protein	fig|6666666.164181.peg.351
63	PF00830	Ribosomal_L28	36.8	Ribosomal L28 family	fig|6666666.164181.peg.1848
64	PF00831	Ribosomal_L29	33.7	Ribosomal L29 protein	fig|6666666.164181.peg.1706

Appendix 2. To be continued

	PFAM ID	HMM name	HMM cutoff	Description of bacterial genes	RS-JS1-cSAG Protein ID in RAST
65	PF00297	Ribosomal_L3	66.35	Ribosomal protein L3	fig|6666666.164181.peg.1044
66	PF01783	Ribosomal_L32p	28.2	Ribosomal L32p protein family	
67	PF01632	Ribosomal_L35p	29.5	Ribosomal protein L35	fig|6666666.164181.peg.48
68	PF00573	Ribosomal_L4	99.95	Ribosomal protein L4/L1 family	fig|6666666.164181.peg.1045
69	PF00281	Ribosomal_L5	34.55	Ribosomal protein L5	fig|6666666.164181.peg.1072
70	PF00673	Ribosomal_L5_C	61.25	ribosomal L5P family C-terminus	fig|6666666.164181.peg.1072
71	PF00347	Ribosomal_L6	53.25	Ribosomal protein L6	fig|6666666.164181.peg.990
72	PF03948	Ribosomal_L9_C	37.95	Ribosomal protein L9, C-terminal domain	fig|6666666.164181.peg.715
73	PF01281	Ribosomal_L9_N	30.55	Ribosomal protein L9, N-terminal domain	fig|6666666.164181.peg.715
74	PF00338	Ribosomal_S10	58.05	Ribosomal protein S10p/S20e	fig|6666666.164181.peg.1043
75	PF00411	Ribosomal_S11	73.45	Ribosomal protein S11	fig|6666666.164181.peg.322
76	PF00164	Ribosomal_S12	84.3	Ribosomal protein S12	fig|6666666.164181.peg.556
77	PF00416	Ribosomal_S13	56.2	Ribosomal protein S13/S18	fig|6666666.164181.peg.321
78	PF00312	Ribosomal_S15	41.7	Ribosomal protein S15	
79	PF00886	Ribosomal_S16	36.25	Ribosomal protein S16	
80	PF00366	Ribosomal_S17	41.3	Ribosomal protein S17	fig|6666666.164181.peg.1705
81	PF01084	Ribosomal_S18	36.3	Ribosomal protein S18	
82	PF00203	Ribosomal_S19	54.9	Ribosomal protein S19	fig|6666666.164181.peg.1048
83	PF00318	Ribosomal_S2	133.15	Ribosomal protein S2	fig|6666666.164181.peg.1924
84	PF01649	Ribosomal_S20p	35.95	Ribosomal protein S20	fig|6666666.164181.peg.1379
85	PF00189	Ribosomal_S3_C	49.5	Ribosomal protein S3, C-terminal domain	fig|6666666.164181.peg.1708
86	PF00163	Ribosomal_S4	32.9	Ribosomal protein S4/S9 N-terminal domain	fig|6666666.164181.peg.323
87	PF00333	Ribosomal_S5	42.55	Ribosomal protein S5, N-terminal domain	fig|6666666.164181.peg.314

Appendix 2. To be continued

	PFAM ID	HMM name	HMM cutoff	Description of bacterial genes	RS-JS1-cSAG Protein ID in RAST
88	PF03719	Ribosomal_S5_C	44.9	Ribosomal protein S5, C-terminal domain	fig|6666666.164181.peg.314
89	PF01250	Ribosomal_S6	45.2	Ribosomal protein S6	fig|6666666.164181.peg.348
90	PF00177	Ribosomal_S7	95.55	Ribosomal protein S7p/S5e	fig|6666666.164181.peg.555
91	PF00410	Ribosomal_S8	72.35	Ribosomal protein S8	fig|6666666.164181.peg.989
92	PF00380	Ribosomal_S9	70.4	Ribosomal protein S9/S16	fig|6666666.164181.peg.1854
93	PF01782	RimM	30.65	RimM N-terminal domain	
94	PF01000	RNA_pol_A_bac	34.3	RNA polymerase Rpb3/RpoA insert domain	fig|6666666.164181.peg.324
95	PF03118	RNA_pol_A_CTD	38.9	Bacterial RNA polymerase, alpha chain C terminal domain	fig|6666666.164181.peg.324
96	PF01193	RNA_pol_L	34.35	RNA polymerase Rpb3/Rpb11 dimerization domain	fig|6666666.164181.peg.324
97	PF04997	RNA_pol_Rpb1_1	157.25	RNA polymerase Rpb1, domain 1	
98	PF00623	RNA_pol_Rpb1_2	79.8	RNA polymerase Rpb1, domain 2	
99	PF04983	RNA_pol_Rpb1_3	41.85	RNA polymerase Rpb1, domain 3	
100	PF05000	RNA_pol_Rpb1_4	24.6	RNA polymerase Rpb1, domain 4	
101	PF04998	RNA_pol_Rpb1_5	119.4	RNA polymerase Rpb1, domain 5	fig|6666666.164181.peg.557
102	PF04563	RNA_pol_Rpb2_1	39.95	RNA polymerase beta subunit	fig|6666666.164181.peg.643
103	PF04561	RNA_pol_Rpb2_2	37.15	RNA polymerase Rpb2, domain 2	fig|6666666.164181.peg.643
104	PF04565	RNA_pol_Rpb2_3	45.6	RNA polymerase Rpb2, domain 3	fig|6666666.164181.peg.485
105	PF10385	RNA_pol_Rpb2_45	36.6	RNA polymerase beta subunit external 1 domain	fig|6666666.164181.peg.1471
106	PF00562	RNA_pol_Rpb2_6	258.45	RNA polymerase Rpb2, domain 6	fig|6666666.164181.peg.1471
107	PF04560	RNA_pol_Rpb2_7	46.45	RNA polymerase Rpb2, domain 7	fig|6666666.164181.peg.1471
108	PF01765	RRF	101.65	Ribosome recycling factor	
109	PF07499	RuvA_C	16.6	RuvA, C-terminal domain	fig|6666666.164181.peg.1108
110	PF01330	RuvA_N	26.85	RuvA N terminal domain	fig|6666666.164181.peg.1108

Appendix 2. To be continued

	PFAM ID	HMM name	HMM cutoff	Description of bacterial genes	RS-JS1-cSAG Protein ID in RAST
111	PF05491	RuvB_C	49.5	Holliday junction DNA helicase ruvB C-terminus	fig|6666666.164181.peg.1121
112	PF02773	S-AdoMet_synt_C	108.9	S-adenosylmethionine synthetase, C-terminal domain	fig|6666666.164181.peg.269
113	PF02772	S-AdoMet_synt_M	77.05	S-adenosylmethionine synthetase, central domain	fig|6666666.164181.peg.269
114	PF00584	SecE	27.5	SecE/Sec61-gamma subunits of protein translocation complex	fig|6666666.164181.peg.1119
115	PF03840	SecG	29.2	Preprotein translocase SecG subunit	fig|6666666.164181.peg.574
116	PF00344	SecY	180.7	SecY translocase	fig|6666666.164181.peg.316
117	PF02403	Seryl_tRNA_N	43.15	Seryl-tRNA synthetase N-terminal domain	
118	PF01668	SmpB	43.4	SmpB protein	fig|6666666.164181.peg.530
119	PF02978	SRP_SPB	52.8	Signal peptide binding domain	
120	PF00763	THF_DHG_CYH	62.2	Tetrahydrofolate dehydrogenase/cyclohydrolase, catalytic domain	fig|6666666.164181.peg.767
121	PF02882	THF_DHG_CYH_C	103.7	Tetrahydrofolate dehydrogenase/cyclohydrolase, NAD(P)-binding domain	fig|6666666.164181.peg.1799
122	PF00121	TIM	140.2	Triosephosphate isomerase	fig|6666666.164181.peg.573
123	PF08275	Toprim_N	58.6	DNA primase catalytic core, N-terminal domain	fig|6666666.164181.peg.2074
124	PF03461	TRCF	40.05	TRCF domain	fig|6666666.164181.peg.1702
125	PF05698	Trigger_C	55.15	Bacterial trigger factor protein (TF) C-terminus	fig|6666666.164181.peg.1487
126	PF05697	Trigger_N	61.8	Bacterial trigger factor protein (TF)	fig|6666666.164181.peg.1487
127	PF01746	tRNA_m1G_MT	80.75	tRNA (guanine-1)-methyltransferase	
128	PF00750	tRNA-synt_1d	120.85	tRNA synthetases class I (R)	
129	PF01409	tRNA-synt_2d	161.15	tRNA synthetases class II core domain (F)	fig|6666666.164181.peg.181
130	PF01509	TruB_N	83.55	TruB family pseudouridylate synthase (N terminal domain)	fig|6666666.164181.peg.1726
131	PF00627	UBA	13.3	UBA/TS-N domain	
132	PF02130	UPF0054	59.35	Uncharacterized protein family UPF0054	fig|6666666.164181.peg.1176
133	PF02367	UPF0079	46.1	Uncharacterized P-loop hydrolase UPF0079	fig|6666666.164181.peg.1430

Appendix 2. To be continued

	PFAM ID	HMM name	HMM cutoff	Description of bacterial genes	RS-JS1-cSAG Protein ID in RAST
134	PF03652	UPF0081	60.75	Uncharacterized protein family (UPF0081)	fig|6666666.164181.peg.29
135	PF12344	UvrB	33.75	Ultraviolet resistance protein B	fig|6666666.164181.peg.130
136	PF08459	UvrC_HhH_N	81.9	UvrC helix-hairpin-helix N-terminal	fig|6666666.164181.peg.373
137	PF10458	Val_tRNA-synt_C	26.5	Valyl tRNA synthetase tRNA binding arm	fig|6666666.164181.peg.1492
138	PF06071	YchF-GTPase_C	61.9	Protein of unknown function (DUF933)	fig|6666666.164181.peg.682
139	PF06689	zf-C4_ClpX	31.7	ClpX C4-type zinc finger	fig|6666666.164181.peg.1489

Appendix 3. Comparison of metabolism among JS1 genomes

Protein ID in RAST	Function	Categories in figure 4-2	AWNT	CDPM	AQYX	SCGC AD-561-N23	ASOY	TA biofilm MG bin	ASOZ	CDPL
fig|6666666.164181.peg.1317	Branched-chain amino acid transport ATP-binding protein LivF (TC	Amino acid and peptide	58.8	29.44	58.8	0	56.84	56.84	31.13	50.83
fig|6666666.164181.peg.1590	Branched-chain amino acid transport ATP-binding protein LivG (TC	Amino acid and peptide	49	31.33	49	0	47.84	46.8	29.71	44.05
fig|6666666.164181.peg.1591	Branched-chain amino acid transport ATP-binding protein LivF (TC	Amino acid and peptide	45.06	28.19	45.06	0	50	50	30.05	47.11
fig|6666666.164181.peg.1295	Branched-chain amino acid transport system permease protein LivM (TC	Amino acid and peptide	52.56	0	52.56	0	53.22	53.22	0	39.69
fig|6666666.164181.peg.1293	Branched-chain amino acid ABC transporter, amino acid-binding protein	Amino acid and peptide	49.21	0	49.21	0	50.92	30.59	0	35.31
fig|6666666.164181.peg.1593	Branched-chain amino acid transport system permease protein LivM (TC	Amino acid and peptide	39.94	0	39.94	0	38.01	43.29	0	38.11
fig|6666666.164181.peg.1592	High-affinity branched-chain amino acid transport system permease	Amino acid and peptide	38.57	0	38.57	0	37.2	44.61	34.25	39.58
fig|6666666.164181.peg.1589	Branched-chain amino acid ABC transporter, amino acid-binding protein	Amino acid and peptide	29.41	0	29.41	0	30.58	39.69	0	33.73
fig|6666666.164181.peg.18	High-affinity branched-chain amino acid transport system permease	Amino acid and peptide	39.54	25.37	39.54	0	**83.83**	49.61	38.78	38.76
fig|6666666.164181.peg.17	Branched-chain amino acid transport system permease protein LivM (TC	Amino acid and peptide	31.56	0	31.56	0	**77.81**	32.44	0	33.23
fig|6666666.164181.peg.15	Branched-chain amino acid transport ATP-binding protein LivF (TC	Amino acid and peptide	55.13	32.38	55.13	0	**79.42**	**79.42**	32.34	53.78
fig|6666666.164181.peg.16	Branched-chain amino acid transport ATP-binding protein LivG (TC	Amino acid and peptide	51	30.43	51	0	**78.6**	**80.61**	29.51	43.25
fig|6666666.164181.peg.1294	High-affinity branched-chain amino acid transport system permease	Amino acid and peptide	58.56	0	58.56	0	**60.96**	**62.3**	0	44.48
fig|6666666.164181.peg.19	Branched-chain amino acid ABC transporter, amino acid-binding protein	Amino acid and peptide	29.64	0	29.64	0	**80.78**	**81.36**	0	27.48
fig|6666666.164181.peg.1297	Branched-chain amino acid transport ATP-binding protein LivF (TC	Amino acid and peptide	**81.62**	33.48	**81.62**	0	**79.49**	55.13	33.16	59.65
fig|6666666.164181.peg.1296	Branched-chain amino acid transport ATP-binding protein LivG (TC	Amino acid and peptide	**66.54**	30.13	**66.54**	0	**69.8**	54.09	33.48	48
fig|6666666.164181.peg.1316	Branched-chain amino acid transport ATP-binding protein LivG (TC	Amino acid and peptide	**64.46**	35.34	**64.46**	0	**65.29**	**78.86**	43.08	55.46
fig|6666666.164181.peg.159	Oligopeptide transport ATP-binding protein OppF (TC 3.A.1.5.1)	Amino acid and peptide	55.17	50.62	55.17	0	50.34	50.88	48.26	47.22
fig|6666666.164181.peg.160	Oligopeptide transport ATP-binding protein OppD (TC 3.A.1.5.1)	Amino acid and peptide	54.26	52.63	54.26	0	48.32	52.47	45.06	49.69
fig|6666666.164181.peg.1762	Dipeptide transport system permease protein DppB (TC 3.A.1.5.2);	Amino acid and peptide	56.74	45.03	56.74	0	0	38.01	30.5	32.85

Appendix 3 To be continued

Protein ID in RAST	Function	Categories in figure 4-2	AWNT	CDPM	AQYX	SCGC AD-561-N23	ASOY	TA biofilm MG bin	ASOZ	CDPL
fig|6666666.164181.peg.1761	Dipeptide transport system permease protein DppC (TC 3.A.1.5.2)	Amino acid and peptide	56.34	49.26	56.34	0	34.54	33.44	41.2	38.06
fig|6666666.164181.peg.161	Dipeptide transport system permease protein DppC (TC 3.A.1.5.2)	Amino acid and peptide	53.33	47.46	53.33	0	46.91	45.07	42.54	45.35
fig|6666666.164181.peg.1764	ABC-type dipeptide transport system, periplasmic component	Amino acid and peptide	49.74	28.39	49.74	0	0	0	0	0
fig|6666666.164181.peg.1763	Dipeptide transport system permease protein DppB (TC 3.A.1.5.2)	Amino acid and peptide	47.79	36.97	47.79	0	0	36.61	34.55	32.48
fig|6666666.164181.peg.600	Dipeptide transport system permease protein DppB (TC 3.A.1.5.2)	Amino acid and peptide	**95.81**	50.53	**95.81**	0	29.05	48.42	36.42	46.28
fig|6666666.164181.peg.598	Oligopeptide transport ATP-binding protein OppD (TC 3.A.1.5.1)	Amino acid and peptide	**95.8**	52.8	**95.8**	0	53.37	50	46.89	49.84
fig|6666666.164181.peg.599	Dipeptide transport system permease protein DppC (TC 3.A.1.5.2)	Amino acid and peptide	**93.38**	44.14	**93.38**	0	43.23	42.91	37.5	43.28
fig|6666666.164181.peg.1753	Dipeptide-binding ABC transporter, periplasmic substrate-binding	Amino acid and peptide	**88.3**	29.36	**88.3**	0	0	30.36	24.81	28.96
fig|6666666.164181.peg.162	Dipeptide transport system permease protein DppB (TC 3.A.1.5.2)	Amino acid and peptide	**88.24**	47.83	**88.24**	0	25.97	48.76	36.67	45.62
fig|6666666.164181.peg.597	Oligopeptide transport ATP-binding protein OppF (TC 3.A.1.5.1)	Amino acid and peptide	**97.24**	51.6	**97.24**	0	50.63	55.09	47.06	47.15
fig|6666666.164181.peg.163	Dipeptide-binding ABC transporter, periplasmic substrate-binding	Amino acid and peptide	**92.1**	**91.67**	**92.1**	0	0	35.51	22.58	29.91
fig|6666666.164181.peg.1447	Dipeptide transport system permease protein DppC (TC 3.A.1.5.2)	Amino acid and peptide	46.27	**94.86**	46.27	0	43.81	**83.74**	41.07	45.36
fig|6666666.164181.peg.1446	Dipeptide transport system permease protein DppB (TC 3.A.1.5.2)	Amino acid and peptide	43.41	**93.11**	43.41	0	25.37	**81.42**	35.93	46.34
fig|6666666.164181.peg.1733	D-aminopeptidase dipeptide-binding protein DppA (EC 3.4.11.-)	Amino acid and peptide	**90.52**	0	**90.52**	0	0	**69.4**	0	**74.35**
fig|6666666.164181.peg.1445	Dipeptide-binding ABC transporter, periplasmic substrate-binding	Amino acid and peptide	**85.16**	**86.21**	**85.16**	0	0	**78.06**	21.63	32.46
fig|6666666.164181.peg.1448	Oligopeptide transport ATP-binding protein OppD (TC 3.A.1.5.1)	Amino acid and peptide	**63.61**	**94.8**	**63.61**	0	**63.35**	**71.91**	49.85	52.38
fig|6666666.164181.peg.790	Oligopeptide transport ATP-binding protein OppD (TC 3.A.1.5.1)	Amino acid and peptide	59.18	50	59.18	0	50	50	36.73	41.82
fig|6666666.164181.peg.1438	Oligopeptide transport system permease protein OppC (TC 3.A.1.5.1)	Amino acid and peptide	40	41.48	40	0	36.43	39.26	41.18	38.81
fig|6666666.164181.peg.656	Oligopeptide transport system permease protein OppB (TC 3.A.1.5.1)	Amino acid and peptide	34.78	31.8	34.78	0	0	50	25.12	35.27
fig|6666666.164181.peg.1572	Dipeptide-binding ABC transporter, periplasmic substrate-binding	Amino acid and peptide	29.65	40.35	29.65	0	0	29.59	0	29.37

Appendix 3 To be continued

Protein ID in RAST	Function	Categories in figure 4-2	AWNT	CDPM	AQYX	SCGC AD-561-N23	ASOY	TA biofilm MG bin	ASOZ	CDPL
fig|6666666.164181.peg.638	Oligopeptide ABC transporter, periplasmic oligopeptide-binding protein	Amino acid and peptide	0	0	0	0	0	37.42	0	0
fig|6666666.164181.peg.964	Oligopeptide transport ATP-binding protein OppD (TC 3.A.1.5.1)	Amino acid and peptide	55.94	55.88	55.94	0	58.02	**60.24**	45.8	47.58
fig|6666666.164181.peg.789	Oligopeptide transport ATP-binding protein OppF (TC 3.A.1.5.1)	Amino acid and peptide	**71.29**	47.85	**71.29**	0	47.67	47.83	43.17	44.04
fig|6666666.164181.peg.385	Dipeptide transport system permease protein DppB (TC 3.A.1.5.2)	Amino acid and peptide	52.78	46.76	52.78	0	0	46.21	42.21	47.76
fig|6666666.164181.peg.386	Oligopeptide transport system permease protein OppC (TC 3.A.1.5.1)	Amino acid and peptide	38.23	38.85	38.23	0	30.23	39.56	37.58	38.63
fig|6666666.164181.peg.330	Oligopeptide ABC transporter, periplasmic oligopeptide-binding protein	Amino acid and peptide	26.34	28.49	26.34	0	0	30.06	27.05	42.59
fig|6666666.164181.peg.1239	Dipeptide transport system permease protein DppC (TC 3.A.1.5.2)	Amino acid and peptide	**96.82**	45.14	**96.82**	0	41.75	42.33	38.99	**96.21**
fig|6666666.164181.peg.1426	Dipeptide-binding ABC transporter, periplasmic substrate-binding	Amino acid and peptide	**95.79**	25.17	**95.79**	0	0	25.54	0	**96.14**
fig|6666666.164181.peg.1240	Dipeptide transport system permease protein DppB (TC 3.A.1.5.2)	Amino acid and peptide	**95.45**	39.03	**95.45**	0	24.51	39.55	34.51	**95.1**
fig|6666666.164181.peg.1238	Oligopeptide transport ATP-binding protein OppD (TC 3.A.1.5.1)	Amino acid and peptide	**94.21**	51.67	**94.21**	0	51.1	51.91	44.1	**93.9**
fig|6666666.164181.peg.1224	Oligopeptide transport ATP-binding protein OppF (TC 3.A.1.5.1)	Amino acid and peptide	**88.67**	57.45	**88.67**	0	50.79	51.31	49.22	**88.73**
fig|6666666.164181.peg.750	Dipeptide transport system permease protein DppB (TC 3.A.1.5.2)	Amino acid and peptide	**91.08**	45.51	**91.08**	0	27.43	40.87	35.67	41.91
fig|6666666.164181.peg.751	Oligopeptide ABC transporter, periplasmic oligopeptide-binding protein	Amino acid and peptide	**90.61**	27.74	**90.61**	0	0	0	0	25.97
fig|6666666.164181.peg.749	Dipeptide transport system permease protein DppC (TC 3.A.1.5.2)	Amino acid and peptide	**90.2**	40.55	**90.2**	0	37.11	39.67	40.07	39.05
fig|6666666.164181.peg.765	Oligopeptide transport ATP-binding protein OppF (TC 3.A.1.5.1)	Amino acid and peptide	**63.38**	**90.41**	**63.38**	0	56.16	56.16	46.48	55.56
fig|6666666.164181.peg.577	Oligopeptide transport ATP-binding protein OppF (TC 3.A.1.5.1)	Amino acid and peptide	**93.52**	**94.44**	**93.52**	0	**61.28**	**64.32**	55.35	52.98
fig|6666666.164181.peg.576	Oligopeptide transport ATP-binding protein OppD (TC 3.A.1.5.1)	Amino acid and peptide	**90.32**	**91.02**	**90.32**	0	**63.53**	59.94	48.07	51.55
fig|6666666.164181.peg.963	Oligopeptide transport ATP-binding protein OppF (TC 3.A.1.5.1)	Amino acid and peptide	**63.86**	**62.31**	**63.86**	0	**60.66**	**61.97**	49.38	48.5
fig|6666666.164181.peg.2116	hypothetical protein	BMC Shell	**86.13**	0	**86.13**	0	0	**69.54**	0	59.2
fig|6666666.164181.peg.26	Formate dehydrogenase alpha subunit (EC 1.2.1.2) @ selenocysteine-	glycolysis	0	35.37	0	0	0	**82.5**	0	0
fig|6666666.164181.peg.28	Formate dehydrogenase chain D (EC 1.2.1.2)	glycolysis	**86.38**	0	**86.38**	0	44.62	44.62	0	0

Appendix 3. To be continued

Protein ID in RAST	Function	Categories in figure 4-2	AWNT	CDPM	AQYX	SCGC AD-561-N23	ASOY	TA biofilm MG bin	ASOZ	CDPL
fig|6666666.164181.peg.622	NAD-dependent formate dehydrogenase alpha subunit @ selenocysteine-	glycolysis	0	27.14	0	0	57.25	**63.29**	0	26.07
fig|6666666.164181.peg.693	Formate dehydrogenase chain D (EC 1.2.1.2)	glycolysis	32.47	0	32.47	0	32.26	32.26	0	0
fig|6666666.164181.peg.79	Fructose-bisphosphate aldolase class II (EC 4.1.2.13)	glycolysis	**94.43**	0	**94.43**	0	**87.46**	**86.67**	0	36.5
fig|6666666.164181.peg.336	Enolase (EC 4.2.1.11)	glycolysis	**94.37**	0	**94.37**	0	0	**88.76**	46.39	0
fig|6666666.164181.peg.572	NAD-dependent glyceraldehyde-3-phosphate dehydrogenase (EC	glycolysis	**93.96**	0	**93.96**	0	**78.25**	**78.25**	0	0
fig|6666666.164181.peg.573	Triosephosphate isomerase (EC 5.3.1.1)	glycolysis	**89.72**	0	**89.72**	0	**72.73**	0	0	0
fig|6666666.164181.peg.718	Pyruvate kinase (EC 2.7.1.40)	glycolysis	**90.93**	**89.66**	**90.93**	**90.08**	**75.37**	**75.37**	0	0
fig|6666666.164181.peg.823	2,3-bisphosphoglycerate-independent phosphoglycerate mutase, archaeal	glycolysis	**92.54**	0	**92.54**	0	**75.37**	**77.58**	0	0
fig|6666666.164181.peg.995	Phosphoglycerate mutase family	glycolysis	**87.44**	**84.19**	**87.44**	0	30.29	0	0	0
fig|6666666.164181.peg.1067	Glucokinase (EC 2.7.1.2)	glycolysis	**91.32**	0	**91.32**	**92.49**	29.97	**69.54**	0	**67.62**
fig|6666666.164181.peg.1156	Phosphoenolpyruvate carboxykinase [GTP] (EC 4.1.1.32)	glycolysis	**82.84**	0	**82.84**	0	0	**64.15**	0	**75.53**
fig|6666666.164181.peg.1410	Phosphomannomutase (EC 5.4.2.8)	glycolysis	0	**90.67**	0	0	0	**60.19**	0	0
fig|6666666.164181.peg.1419	Phosphomannomutase (EC 5.4.2.8)	glycolysis	0	0	0	0	0	**64.35**	0	0
fig|6666666.164181.peg.1756	Glucose-6-phosphate isomerase, archaeal (EC 5.3.1.9)	glycolysis	51.85	0	51.85	0	0	0	0	0
fig|6666666.164181.peg.1902	6-phosphofructokinase (EC 2.7.1.11)	glycolysis	**92.52**	**89.47**	**92.52**	0	**75.24**	37.9	0	0
fig|6666666.164181.peg.1904	NAD(P)-dependent glyceraldehyde 3-phosphate dehydrogenase archaeal	glycolysis	0	0	0	0	0	0	0	0
fig|6666666.164181.peg.1905	Phosphoglycerate kinase (EC 2.7.2.3)	glycolysis	36.88	0	36.88	0	34.86	47.92	0	0
fig|6666666.164181.peg.83	Pyruvate,phosphate dikinase (EC 2.7.9.1)	glycolysis	**87.8**	0	**87.8**	0	0	29.52	31.9	0
fig|6666666.164181.peg.623	Dihydrolipoamide dehydrogenase (EC 1.8.1.4)	glycolysis	**91.3**	0	**91.3**	0	0	0	0	0
fig|6666666.164181.peg.1029	Pyruvate-flavodoxin oxidoreductase (EC 1.2.7.-)	glycolysis	0	0	0	26.9	**79.43**	0	29.82	0
fig|6666666.164181.peg.1073	Acetyl-CoA hydrolase	glycolysis	**90.8**	0	**90.8**	0	0	0	0	0

Appendix 3. To be continued

Protein ID in RAST	Function	Categories in figure 4-2	AWNT	CDPM	AQYX	SCGC AD-561-N23	ASOY	TA biofilm MG bin	ASOZ	CDPL
fig|6666666.164181.peg.1084	Dihydrolipoamide acetyltransferase component of pyruvate	glycolysis	39.42	**89.77**	39.42	0	0	39.21	0	32.42
fig|6666666.164181.peg.1091	Pyruvate formate-lyase (EC 2.3.1.54)	glycolysis	**92.21**	29.56	**92.21**	0	**76.08**	**76.08**	0	**79.14**
fig|6666666.164181.peg.1093	Alcohol dehydrogenase (EC 1.1.1.1)	glycolysis	**89.49**	35.43	**89.49**	33.62	0	40.16	0	36.41
fig|6666666.164181.peg.1349	Fumarate hydratase class I, aerobic (EC 4.2.1.2); L(+)-tartrate	glycolysis	**83.15**	0	**83.15**	0	0	**77**	0	**67.39**
fig|6666666.164181.peg.1350	Fumarate hydratase class I, aerobic (EC 4.2.1.2); L(+)-tartrate	glycolysis	**88.79**	0	**88.79**	0	0	0	0	**69.4**
fig|6666666.164181.peg.1467	D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95)	glycolysis	46.48	44	46.48	0	**64.14**	57.32	31.25	36.51
fig|6666666.164181.peg.1641	CoA-acylating propionaldehyde dehydrogenase # propanediol utilization	glycolysis	**94.67**	**95**	**94.67**	0	**90.67**	0	0	**82.67**
fig|6666666.164181.peg.1655	NADP-dependent malic enzyme (EC 1.1.1.40)	glycolysis	44.69	0	44.69	0	0	**77.97**	0	47.32
fig|6666666.164181.peg.1752	Pyruvate,phosphate dikinase (EC 2.7.9.1)	glycolysis	**89.34**	0	**89.34**	0	0	0	0	0
fig|6666666.164181.peg.2024	Biotin carboxylase of acetyl-CoA carboxylase (EC 6.3.4.14) / Biotin	glycolysis	**85.4**	0	**85.4**	**77.37**	59.42	0	0	0
fig|6666666.164181.peg.2025	Oxaloacetate decarboxylase beta chain (EC 4.1.1.3)	glycolysis	**95.42**	0	**91.13**	**72.76**	**85.75**	0	0	0
fig|6666666.164181.peg.2140	Pyruvate dehydrogenase E1 component alpha subunit (EC 1.2.4.1)	glycolysis	36.23	0	36.23	0	0	41.12	0	40.88
fig|6666666.164181.peg.222	ABC-type tungstate transport system, ATP-binding protein	Mineral and organic ion	**83.78**	39.01	**83.78**	0	**63.13**	**63.13**	35.27	36.87
fig|6666666.164181.peg.223	ABC-type tungstate transport system, permease protein	Mineral and organic ion	**95.65**	**95.03**	**95.65**	0	**72.17**	**72.61**	0	27.98
fig|6666666.164181.peg.224	ABC-type tungstate transport system, periplasmic binding protein	Mineral and organic ion	**70.22**	**88.24**	**70.22**	0	**66.3**	**66.3**	0	0
fig|6666666.164181.peg.1607	Inositol transport system ATP-binding protein	Monosaccharide	41.74	37.02	41.74	0	49.79	49.79	40.09	41.74
fig|6666666.164181.peg.1691	Myo-inositol 2-dehydrogenase 1 (EC 1.1.1.18)	Monosaccharide	35	24.84	35	28.48	28.3	34.18	0	34.44
fig|6666666.164181.peg.1085	Inositol transport system permease protein	Monosaccharide	31.65	32.72	31.65	0	30.37	35.29	45.31	28.85
fig|6666666.164181.peg.146	Inositol transport system ATP-binding protein	Monosaccharide	40.38	36.56	40.38	0	**82.95**	**82.95**	37.86	40.38
fig|6666666.164181.peg.147	Inositol transport system ATP-binding protein	Monosaccharide	43.69	46.55	43.69	0	**73.98**	**73.98**	40.71	45.19
fig|6666666.164181.peg.148	Inositol transport system permease protein	Monosaccharide	35.07	33.68	35.07	0	**79.39**	**83.56**	28.43	38.58

Appendix 3. To be continued

Protein ID in RAST	Function	Categories in figure 4-2	AWNT	CDPM	AQYX	SCGC AD-561-N23	ASOY	TA biofilm MG bin	ASOZ	CDPL
fig|6666666.164181.peg.149	Inositol transport system sugar-binding protein	Monosaccharide	30.28	30.28	30.28	29.9	**81.56**	**81.56**	0	0
fig|6666666.164181.peg.1228	Myo-inositol 2-dehydrogenase (EC 1.1.1.18)	Monosaccharide	**85.03**	0	**90.2**	0	0	26.89	0	**90.42**
fig|6666666.164181.peg.833	Inosose isomerase (EC 5.3.99.11)	Monosaccharide	0	0	0	0	**61.23**	**61.23**	0	**66.18**
fig|6666666.164181.peg.143	Myo-inositol 2-dehydrogenase 1 (EC 1.1.1.18)	Monosaccharide	**83.78**	26.91	**83.78**	29.52	28.45	**81.65**	0	**74.17**
fig|6666666.164181.peg.1564	Ribose ABC transport system, ATP-binding protein RbsA (TC 3.A.1.2.1)	Monosaccharide	40.87	41.56	40.87	0	37.55	37.96	32.46	51.22
fig|6666666.164181.peg.1383	Ribose ABC transport system, ATP-binding protein RbsA (TC 3.A.1.2.1)	Monosaccharide	44.27	43.06	44.27	0	36.35	45.1	38.38	45.05
fig|6666666.164181.peg.1565	Ribose ABC transport system, periplasmic ribose-binding protein RbsB	Monosaccharide	38.46	36.36	38.46	0	30.43	30.43	0	41.08
fig|6666666.164181.peg.1384	Ribose ABC transport system, permease protein RbsC (TC 3.A.1.2.1)	Monosaccharide	40.93	40.21	40.93	0	34.34	37.78	28.21	37.68
fig|6666666.164181.peg.1385	Ribose ABC transport system, high affinity permease RbsD (TC	Monosaccharide	35.16	35.94	35.16	0	0	0	0	0
fig|6666666.164181.peg.1563	Ribose ABC transport system, permease protein RbsC (TC 3.A.1.2.1)	Monosaccharide	31.29	33.23	31.29	0	28	39.84	0	41.82
fig|6666666.164181.peg.673	Ribose ABC transport system, ATP-binding protein RbsA (TC 3.A.1.2.1)	Monosaccharide	41.6	40.24	41.6	0	36.18	39.63	34.14	47.71
fig|6666666.164181.peg.674	Ribose ABC transport system, permease protein RbsC (TC 3.A.1.2.1)	Monosaccharide	39.09	39.29	39.09	0	33.88	35.38	31.94	38.51
fig|6666666.164181.peg.1858	Ribose ABC transport system, permease protein RbsC (TC 3.A.1.2.1)	Monosaccharide	34.69	35.94	34.69	0	35.13	**83.45**	0	36.89
fig|6666666.164181.peg.1864	Ribose ABC transport system, permease protein RbsC (TC 3.A.1.2.1)	Monosaccharide	34.36	35.05	34.36	0	37.41	**80**	32.92	40.59
fig|6666666.164181.peg.1859	Ribose ABC transport system, permease protein RbsC (TC 3.A.1.2.1)	Monosaccharide	33.81	34.29	33.81	0	36.89	**79.55**	25.53	39.6
fig|6666666.164181.peg.1860	ribose ABC transporter, periplasmic binding protein	Monosaccharide	28.46	28.84	28.46	33.71	31.97	**90.75**	0	29.38
fig|6666666.164181.peg.1538	Ribokinase (EC 2.7.1.15)	Monosaccharide	30.23	29.41	30.23	29.3	**78.01**	**78.01**	0	0
fig|6666666.164181.peg.313	L-ribulose-5-phosphate 4-epimerase (EC 5.1.3.4)	Monosaccharide	0	33.65	0	0	33.97	**75.12**	0	0
fig|6666666.164181.peg.327	L-ribulose-5-phosphate 4-epimerase (EC 5.1.3.4)	Monosaccharide	0	33.65	0	0	33.97	**75.12**	0	0
fig|6666666.164181.peg.2095	Ribose ABC transport system, permease protein RbsC (TC 3.A.1.2.1)	Monosaccharide	**97.42**	**95.48**	**97.42**	0	40.88	42.86	27.13	43.23
fig|6666666.164181.peg.2099	Ribokinase (EC 2.7.1.15)	Monosaccharide	**84.72**	37.84	**84.72**	0	30.11	30.11	0	0

Appendix 3. To be continued

Protein ID in RAST	Function	Categories in figure 4-2	AWNT	CDPM	AQYX	SCGC AD-561-N23	ASOY	TA biofilm MG bin	ASOZ	CDPL
fig|6666666.164181.peg.1386	Ribulose-5-phosphate 4-epimerase and related epimerases and aldolases	Monosaccharide	0	**90.83**	0	0	**68.87**	**68.87**	0	29.82
fig|6666666.164181.peg.2096	Ribose ABC transport system, ATP-binding protein RbsA (TC 3.A.1.2.1)	Monosaccharide	**94.63**	**90.85**	**94.63**	0	37.96	42.27	36.08	52.78
fig|6666666.164181.peg.993	Ribose ABC transport system, periplasmic ribose-binding protein RbsB	Monosaccharide	**91.22**	**91.22**	**91.22**	**95**	29.52	30.3	0	32
fig|6666666.164181.peg.2097	Ribose ABC transport system, high affinity permease RbsD (TC	Monosaccharide	**88.79**	**88.79**	**88.79**	0	0	0	0	0
fig|6666666.164181.peg.1535	Ribose 5-phosphate isomerase B (EC 5.3.1.6)	Monosaccharide	**94.54**	**95.08**	**94.54**	0	46.15	**77.47**	0	**79.23**
fig|6666666.164181.peg.1286	Transketolase, N-terminal section (EC 2.2.1.1)	Monosaccharide	**91.69**	**91.3**	**91.69**	0	**78.04**	**75.5**	0	**88.58**
fig|6666666.164181.peg.1133	Transketolase (EC 2.2.1.1)	Monosaccharide	**90.59**	**90.94**	**90.59**	0	29.03	**73.71**	0	**66.67**
fig|6666666.164181.peg.1457	Transketolase (EC 2.2.1.1)	Monosaccharide	**86.14**	**86.14**	**86.14**	0	39.24	**83.17**	0	**81**
fig|6666666.164181.peg.2058	Xylulose kinase (EC 2.7.1.17)	Monosaccharide	26.97	34.51	26.97	0	24.56	24.56	41.65	34.34
fig|6666666.164181.peg.1695	Xylose isomerase (EC 5.3.1.5)	Monosaccharide	0	0	0	0	0	**82.32**	0	0
fig|6666666.164181.peg.1634	Xylulose kinase (EC 2.7.1.17)	Monosaccharide	24.07	**89.75**	24.07	0	29.07	29.07	31.45	**85.75**
fig|6666666.164181.peg.311	predicted xylose isomerase	Monosaccharide	0	0	0	0	0	**78.26**	0	0
fig|6666666.164181.peg.329	predicted xylose isomerase	Monosaccharide	0	0	0	0	0	**78.26**	0	0
fig|6666666.164181.peg.97	Alpha-glucoside transport system permease protein AglG	Oligosaccharide	36.7	30.8	36.7	0	**82.46**	31.63	27.19	28.46
fig|6666666.164181.peg.96	ABC alpha-glucoside transporter, inner membrane subunit AglF	Oligosaccharide	35.29	35.29	35.29	0	**86.41**	55.81	31.75	28.14
fig|6666666.164181.peg.95	Alpha-glucosides-binding periplasmic protein AglE precursor	Oligosaccharide	26.9	27.35	26.9	0	**79.76**	0	0	0
fig|6666666.164181.peg.249	Fructokinase (EC 2.7.1.4)	Oligosaccharide	23.87	**75.84**	23.87	0	25	25	0	0
fig|6666666.164181.peg.1495	Fructokinase (EC 2.7.1.4)	Oligosaccharide	0	**91.04**	0	0	27.36	27.36	0	0
fig|6666666.164181.peg.1997	Fructokinase (EC 2.7.1.4)	Oligosaccharide	0	**91.36**	0	0	28.74	28.74	0	0
fig|6666666.164181.peg.1148	Beta-galactosidase (EC 3.2.1.23)	Oligosaccharide	0	0	0	0	0	0	0	0
fig|6666666.164181.peg.1903	Galactose-1-phosphate uridylyltransferase (EC 2.7.7.10)	Oligosaccharide	**90.45**	29.82	**90.45**	0	52.06	0	0	0

Appendix 3. To be continued

Protein ID in RAST	Function	Categories in figure 4-2	AWNT	CDPM	AQYX	SCGC AD-561-N23	ASOY	TA biofilm MG bin	ASOZ	CDPL
fig|6666666.164181.peg.1711	Galactose-1-phosphate uridylyltransferase (EC 2.7.7.10)	Oligosaccharide	**89.55**	30.18	**89.55**	0	51.76	0	0	0
fig|6666666.164181.peg.957	Galactokinase (EC 2.7.1.6)	Oligosaccharide	**87.47**	**85.93**	**87.47**	0	0	0	0	0
fig|6666666.164181.peg.839	Predicted galactoside ABC transporter type II, sugar-binding protein	Oligosaccharide	27.65	27.02	27.65	0	29.62	0	22.35	23.55
fig|6666666.164181.peg.840	Predicted galactoside ABC transporter type II, permease protein 1	Oligosaccharide	34.04	34.46	34.04	0	30.68	32.26	29.97	27.23
fig|6666666.164181.peg.1149	Predicted galactoside ABC transporter type II, permease protein 2	Oligosaccharide	31.14	32.96	31.14	0	40.62	34.6	31.32	29.63
fig|6666666.164181.peg.400	Multiple sugar ABC transporter, ATP-binding protein	Oligosaccharide	**88.59**	49.48	**88.59**	0	**80.05**	**80.05**	39.67	48.99
fig|6666666.164181.peg.1543	ABC sugar transporter, ATP-binding component	Oligosaccharide	42.31	42.24	42.31	0	42.36	38.61	41.84	38.82
fig|6666666.164181.peg.1542	Maltose/maltodextrin transport ATP-binding protein MalK (EC 3.6.3.19)	Oligosaccharide	36.89	41.59	36.89	0	40.57	37.54	43.72	38.99
fig|6666666.164181.peg.1539	Maltose/maltodextrin ABC transporter, permease protein MalG	Oligosaccharide	33.2	39.31	33.2	0	36.76	32.86	43.75	34.18
fig|6666666.164181.peg.1541	ABC transport system, sugar-binding protein	Oligosaccharide	34.69	24.17	34.69	0	25.91	0	26.34	0
fig|6666666.164181.peg.1540	Dihydroxyacetone ABC transport system, permease protein 2 # predicted	Oligosaccharide	34.27	35.25	34.27	0	31.05	40.23	33.73	29.7
fig|6666666.164181.peg.353	Glucose-6-phosphate isomerase, archaeal II (EC 5.3.1.9) / Mannose-6-	Oligosaccharide	0	0	0	0	0	0	0	0
fig|6666666.164181.peg.1089	Alcohol dehydrogenase (EC 1.1.1.1)	Oligosaccharide	**93.69**	31.43	**93.69**	0	31.36	31.68	29.9	29.34
fig|6666666.164181.peg.1181	Putative sucrose phosphorylase (EC 2.4.1.7)	Oligosaccharide	27.04	**81.56**	27.04	0	0	57.25	0	**61.56**
fig|6666666.164181.peg.1209	Putative sucrose phosphorylase (EC 2.4.1.7)	Oligosaccharide	0	**86.02**	0	0	0	53.23	0	**61.08**
fig|6666666.164181.peg.33	Acetyl-CoA synthetase (ADP-forming) alpha and beta chains, putative	Wood Ljungdahl pathway	**76.99**	28.76	**76.99**	0	0	**73.45**	0	**70.35**
fig|6666666.164181.peg.34	Acetyl-CoA synthetase (ADP-forming) alpha and beta chains, putative	Wood Ljungdahl pathway	**88.25**	26.8	**88.25**	0	27.96	**67.65**	0	**83.18**
fig|6666666.164181.peg.1377	Formate--tetrahydrofolate ligase (EC 6.3.4.3)	Wood Ljungdahl pathway	**95.77**	0	**95.77**	0	36.9	0	0	**89.89**
fig|6666666.164181.peg.437	[NiFe] hydrogenase nickel incorporation protein HypA	Wood Ljungdahl pathway	**88.5**	**82.3**	**88.5**	0	57.52	57.52	0	0
fig|6666666.164181.peg.1608	[NiFe] hydrogenase metallocenter assembly protein HypF	Wood Ljungdahl pathway	**87.78**	0	**87.78**	0	0	**67.56**	0	0
fig|6666666.164181.peg.1609	[NiFe] hydrogenase metallocenter assembly protein HypC	Wood Ljungdahl pathway	**88.73**	0	**88.73**	0	0	**70.59**	0	0

Appendix 3. To be continued

Protein ID in RAST	Function	Categories in figure 4-2	AWNT	CDPM	AQYX	SCGC AD-561-N23	ASOY	TA biofilm MG bin	ASOZ	CDPL
fig|6666666.164181.peg.1610	[NiFe] hydrogenase metallocenter assembly protein HypD	Wood Ljungdahl pathway	0	0	0	0	0	**72.55**	0	0
fig|6666666.164181.peg.1668	[NiFe] hydrogenase metallocenter assembly protein HypE	Wood Ljungdahl pathway	**88.83**	0	**88.83**	0	0	**75.53**	0	0
fig|6666666.164181.peg.1696	[NiFe] hydrogenase nickel incorporation-associated protein HypB	Wood Ljungdahl pathway	**92.9**	0	**92.9**	0	**71.43**	**71.43**	0	0
fig|6666666.164181.peg.880	Xanthine dehydrogenase, molybdenum binding subunit (EC 1.17.1.4)	Wood Ljungdahl pathway	**92.95**	0	**92.95**	0	**79.31**	**86.32**	0	0
fig|6666666.164181.peg.881	Xanthine dehydrogenase, FAD binding subunit (EC 1.17.1.4)	Wood Ljungdahl pathway	**92.47**	0	**92.47**	0	**64.93**	**64.93**	33.33	0
fig|6666666.164181.peg.882	Xanthine dehydrogenase iron-sulfur subunit (EC 1.17.1.4)	Wood Ljungdahl pathway	**93.38**	0	**93.38**	0	**76.97**	**76.97**	48.67	0
fig|6666666.164181.peg.958	hypothetical protein	Wood Ljungdahl pathway	31.45	**91.63**	31.45	0	**74.01**	**72.77**	0	0
fig|6666666.164181.peg.961	5-methyltetrahydrofolate--homocysteine methyltransferase (EC 2.1.1.13)	Wood Ljungdahl pathway	59.4	**96.55**	59.4	0	**78.89**	**78.5**	53.01	0
fig|6666666.164181.peg.982	5-methyltetrahydrofolate--homocysteine methyltransferase (EC 2.1.1.13)	Wood Ljungdahl pathway	59.33	**92.86**	59.33	0	**77.99**	**77.99**	53.33	**74.76**
fig|6666666.164181.peg.1799	Methylenetetrahydrofolate dehydrogenase (NADP+) (EC 1.5.1.5) /	Wood Ljungdahl pathway	0	**87.37**	0	0	**75.09**	**75.09**	0	0

국문초록(ABSTRACT IN KOREAN)

원핵생물은 해양 생태계에서 생지화학적 순환과 유기물의 광물화 등의 중요한 생태학적 역할을 담당한다. 남극해는 이산화탄소를 흡수하고, 차가운 남극 저층수의 형성을 통해 지구 온난화에 반하는 역할을 하므로, 전지구적인 중요성을 지닌다. 또한 남극해는 넓은 범위에 걸쳐 있는 해양으로, 남극해에서 미생물에 의한 탄소 고정과 영양분 순환과 같은 생태적 기능은 전지구적 생태계와 직접적으로 연결되어 있어서 해당 지역에서 미생물의 역할을 이해하는 것은 매우 중요하다. 그러나, 남극해 생태계의 전지구적인 중요성에도 불구하고, 해당 지역에 대한 접근의 어려움 때문에 열대지역이나 온대지역에서의 세균 다양성 또는 미생물 생태 연구에 비해 이 지역에서의 연구가 미미한 실정이다. 또한, 대부분의 다양성 연구가 전통적 분자기법을 사용하여 수행되어 전체 다양성을 깊이 있게 이해하기에 한계가 있어왔다.

따라서, 본 연구에서는 남극해 해수와 퇴적물내 세균 다양성에 대한 이해를 넓히기 위해 16S rRNA 유전자의 염기서열을 대용량 염기서열 분석기법을 사용하여 세균 다양성을 깊이 있게 분석하였다. 또한 세균 다양성 구조는 환경인자에 의해 크게 영향을 받기 때문에 각 세균군의 변화 양상에 대한 이해를 높이기 위해 세균 다양성과 환경인자와의 관계를 분석하였다. 또한, 세균의 생태적인 기능을 이해하기 위해, 해양 퇴적물의 배양을 통해 확보한 세균의 생리특성을 연구하였다. 마지막으로, 배양체가 획득되지 않았으나 혐기 상태의 퇴적물에 우점하는 *Atribacteria* JS1의 대사 기작과 생태적 기능을 이해하기 위해 단일세포분리를 통해 확보한 유전체를 분석하였다.

세균 다양성은 해수와 퇴적물에서 대규모 및 소규모로 구분하여 연구를 수행하였다. 대규모 연구에서는 남극해의 1600에서 3000 km에 이르는 구간

에서 전선의 존재에 따른 표층수의 세균 군집 구조를 연구하였고, 소규모 연구에서는 로스해 해수 및 퇴적물내 세균 군집의 수직적 및 수평적인 분포를 분석하였다. 로스해의 테라노바만의 다양한 지역에서 수층의 깊이에 따른 세균 군집 구조와 주변 환경에 영향을 많이 받는 표층수에서의 세균 군집이 육상으로부터의 거리에 따라 어떤 분포를 보이는지 분석하였다. 퇴적물의 경우, 공간에 따른 해양 퇴적물의 세균 다양성 분포를 분석하기 위해 표층 및 심층 퇴적물의 세균 다양성을 로스해 및 남극 중앙해령에서 채취된 시료를 대상으로 연구 하였다.

남극해 해수 세균 다양성에 대한 대규모 연구에서는 *Bacteroidetes, Alphaproteobacteria,* 및 *Gammaproteobacteria* 가 우점하였다. 이들 세 분류군의 우점과 관계없이, *Cyanobacteria*와 *Actinobacteria*는 Subantarctic Front의 남쪽 지역에서는 거의 관찰되지 않았고, *Verrucomicrobia*와 SAR406는 Polar Front의 남쪽 지역에서 거의 관찰되지 않았는데, 이는 특정 세균군의 분포가 전선에 따라 다르게 영향을 받고 있음을 의미한다. 전선에 따라 뚜렷한 세균 군집의 분포 차이가 관찰되었는데, 이는 각 전선 사이의 구역에서 세균의 대사학적 특성과 기능이 서로 다음을 나타낸다. 서로 높은 상관관계를 가진 환경인자들 중, 온도가 이들 세균 군집의 구분에 가장 큰 영향을 미치는 것으로 나타났다.

로스해 해수의 세균성 다양성에 대한 소규모 연구에서는 기존에 전통적인 분자 기법을 사용했을 때는 나타나지 않았던 다양한 분류군을 확인할 수 있었다. *Bacteroidetes, Alphaproteobacteria,* 및 *Gammaproteobacteria,* 가 수층내에서 우점하였다. 깊이에 따라 *표층수에서는 거의 관찰되지 않았던 Deltaproteobacteria,* SAR406*, Verrucomicrobia, Planctomycetes, Lentisphaerae, Chloroflexi, Gemmatimonadetes, Planctomycetes,* 및 *Firmicutes* 의 비율이 증가하였

고, 세균 다양성 역시 증가하였다. 세균 군집의 수평적인 분포에서는 *Bacteroidetes*의 비율이 연안에 가까울수록 높은 것을 확인할 수 있었는데 이는 조류의 증가와 관련이 있을 것으로 생각된다. 세균 군집 구조는 깊이에 따라 뚜렷한 성층화가 관찰되었고, 세균 군집의 수평적인 차이는 수직적 차이에 비해서는 차이가 덜 한 것으로 확인되었다. 이러한 차이는 수평적인 차이에 비해 깊이에 따라 급격하게 변하는 환경인자로 설명될 수 있었고, 다양한 환경인자 중, 산소 농도 및 질산염과 아질산염의 농도가 세균 군집의 차이를 가장 잘 설명하는 것으로 나타났다.

해양 퇴적물내 세균 다양성의 공간적인 분포 연구에서는 기존에 이 지역에서 관찰되었던 분류군에 비해 다양한 분류군의 존재와 현재까지 배양되지 않은 그룹의 높은 비율을 확인하였다. 퇴적물 세균 군집은 산소의 유무에 따라 뚜렷한 차이를 나타내었는데, 호기 상태의 표층 퇴적물에서는 *Gammaproteobacteria*, *Deltaproteobacteria*, *Bacteroidetes*, 및*Planctomycetes* 가 우점한 반면, 혐기 상태의 퇴적물에서는 후보문 *Atribacteria* JS1, *Chloroflexi*, 및 *Actinobacteria* 가 높은 비율을 차지하는 것으로 나타났다. 후보문 비율 역시 표층보다 혐기 상태의 퇴적물에서 높은 것으로 나타났는데, 이는 퇴적물에서 산화환원전위가 군집 구조 형성에 중요한 역할을 하기 때문으로 생각된다.

세균의 생태적 특성을 살펴보기 위해, 다양성 결과를 기반으로, 기존에 배양되지 않은 세균 그룹이 많은 퇴적물 시료를 선정하였고, 이를 빈영양 배지를 사용하여 낮은 온도에서 배양하였다. 배양의 한계에도 불구하고, 16S rRNA 유전자의 염기서열이 기존에 알려진 균주의 염기서열에 비해 98.65%미만의 유사도를 보이는 균주를 확보하였고, 이들의 영양분 순환과 관련된 기능을 세포외분비 효소 활성 존재 유무로 분석하였다. 그 결과, 확보한 균주의

46%와 25%에 해당하는 균주가 단백질 분해효소 및 지질 분해효소 활성을 나타냈고, 이는 이들 세균이 주요한 유기물 성분의 가수분해에 기여하여 로스해의 낮은 온도에서 탄소 및 질소 순환에 관련된 기능을 수행하고 있을 것으로 생각되어진다.

혐기적 퇴적물에서 풍부하게 관찰되어 온 후보문 *Atribacteria* JS1에 속하는 세균의 생태학적인 기능은 현재까지 배양체가 획득되지 않아 거의 연구가 되지 않은 실정이다. 세균 다양성 분석 결과, 로스해의 혐기 퇴적물에 후보문 *Atribacteria* JS1이 풍부하게 존재하는 것으로 확인이 되어, 이들의 대사기작 및 생태적 기능을 이해하기 위해 단일세포 분리 및 유전체 중폭을 통하여 이들 세균 그룹의 유전체를 획득하고자 하였다. 로스해에 풍부하게 존재하는 단일 종의 유전체가 성공적으로 획득되었고, 높은 커버리지의 *Atribacteria* JS1을 확보하였고, 유전체 분석은 해당 종이 혐기성 메탄 생성 환경에 어떻게 풍부하게 존재할 수 있는지를 설명하였다. 이들은 다양한 종속영양성 기질로부터 아세테이트를 생성하거나 다른 메탄 생성균과 공생을 통해 아세테이트를 산화하는 대사기작는 다양한 대사 기작을 통해 자연계에서 생존 가능성을 높임으로써 해당 서식처에서 우점하고 있는 것으로 생각된다. *Atribacteria* JS1 세균의 유전체 분석을 통해 이들 세균이 로스해에서 탄소 순환과 관련된 생태적 기능에 중요한 역할을 수행함을 밝혔다고 생각된다.

주요어: 해양 세균, 세균 다양성, pyrosequencing, 배양, 단독세포 유전체학, 세포외효소 프로파일, 후보문, *Atribacteria* JS1, 로스해, 남극해